A Short Course in Combinatorial Designs

Ian Anderson
Department of Mathematics
University of Glasgow
Glasgow G12 8QW, UK
e-mail: ia@maths.gla.ac.uk

Iiro Honkala
Department of Mathematics
University of Turku
20014 Turku, Finland
e-mail: honkala@utu.fi

Internet Edition, Spring 1997

This edition is the file `http://www.utu.fi/~honkala/designs.ps`. It may be taken at no charge by all interested persons. Comments and corrections are welcome.

Contents

Chapter 1

Systems of distinct representatives

1.1 Hall's theorem

Example 1.1 Assume that there are five vacant jobs $1, 2, 3, 4$ and 5. Denote by S_i the set of applicants for the job i. If we assume that

$$S_1 = \{A, B, C\}, S_2 = \{D, E\}, S_3 = \{A, D\}, S_4 = \{E\}, S_5 = \{A, E\},$$

is it possible to assign a different person to each job? Clearly, in this specific case the answer is no. Indeed, we must assign E to 4, D to 2, A to 3 and there is nobody left for the job 5. $\square$

Consider the general case.

Definition 1.2 *A system of distinct representatives for the sets S_1, S_2, ..., S_n is an n-tuple $(x_1, x_2, ..., x_n)$ where the elements x_i are distinct and $x_i \in S_i$ for all $i = 1, 2, \ldots, n$.*

In the previous example the sets S_i did not have a system of distinct representatives for the following very simple reason: the union of certain four sets had less than four elements. Clearly, if the sets $S_1, S_2, \ldots, S_n$ have a system of distinct representatives, then the union of any k $(1 \leq k \leq n)$ sets has at least k elements. Surprisingly, this obvious necessary condition is also sufficient.

Theorem 1.3 *(Philip Hall) The sets $S_1, S_2, \ldots, S_n$ have a system of distinct representatives if and only if*

for every $k = 1, 2, ..., n$ the union of any k sets has at least k elements.

Proof. We prove the sufficiency of the condition by induction on n. The case $n = 1$ is clear. Assume that the claim holds for any collection with less than n sets.

Case 1: *For each k, $1 \leq k < n$, the union of any k sets contains more than k elements.* Take any of the sets, and choose any of its elements x as its representative, and remove x from all the other sets. The union of any $s \leq n-1$ of the remaining $n-1$ sets has at least s elements, and therefore the remaining sets have a system of distinct representatives, which together with x gives a system of distinct representatives for the original family.

Case 2: *The union of some k, $1 \leq k < n$ of the sets contains exactly k elements.* By the induction hypothesis, these k sets have a system of distinct representatives. Remove these k elements from the remaining $n-k$ sets. Take any s of these remaining sets. Their union contains at least s elements; otherwise the union of these s sets and the k sets has less than $s+k$ elements. Consequently, also these remaining $n-k$ sets have a system of distinct representatives by the induction hypothesis. Together these two systems of distinct representatives give a system of distinct representatives for the original family. $\square$

1.2 Latin squares

Definition 1.4 *An $n \times n$ array (matrix) is a Latin square of order n, if each of the numbers $1, 2, \ldots, n$ occurs once in each row and each column.*

More generally, we could of course use any n different symbols instead of $1, 2, \ldots, n$.

Definition 1.5 *An $r \times n$ array is called a Latin rectangle, if each of the numbers $1, 2, \ldots n$ occurs once in each row and at most once in each column.*

The following theorem shows that we can build Latin squares one row at a time.

Theorem 1.6 *If $r < n$, any given $r \times n$ Latin rectangle can be extended to an $(r+1) \times n$ Latin rectangle.*

Proof. Define

$$S_j = \text{the set of integers not yet occuring in the } j\text{-th column.}$$

It is sufficient to prove that the sets $S_1, S_2, \ldots, S_n$ have a system of distinct representatives. Let $A_1, A_2, \ldots, A_k$ be any k of the sets $S_1, S_2, \ldots, S_n$. There are $n - r$ elements in each set, and therefore $|A_1| + |A_2| + \ldots |A_k| = k(n-r)$.

Each number occurs once in each of the r rows and hence in $n-r$ of the sets S_i, and consequently in *at most* $n-r$ of the sets A_i. Therefore $|A_1 \cup \ldots \cup A_k| \geq k(n-r)/(n-r) = k$. The claim now follows from Hall's theorem. □

1.3 Optimal assignment problem

In Example 1.1 we did not pay any attention to *how suitable* the individual applicants were to each job.

Example 1.7 Four persons a_1, a_2, a_3 and a_4 should be assigned to the jobs J_1, J_2, J_3 and J_4, and the left-hand side table below gives the available information about their suitability:

$$\begin{array}{ccccc} & J_1 & J_2 & J_3 & J_4 \\ a_1 & 11 & 9 & 1 & 4 \\ a_2 & 8 & 13 & 7 & 6 \\ a_3 & 12 & 2 & 14 & 11 \\ a_4 & 10 & 13 & 15 & 2 \end{array}, \qquad \begin{array}{ccccc} & J_1 & J_2 & J_3 & J_4 \\ a_1 & -11 & -9 & -1 & -4 \\ a_2 & -8 & -13 & -7 & -6 \\ a_3 & -12 & -2 & -14 & -11 \\ a_4 & -10 & -13 & -15 & -2 \end{array}.$$

Our goal is to choose four entries, no two in the same row or column, in such a way that the sum of the entries is maximized. For practical reasons we instead consider the *unsuitability* of the applicants and try to *minimize* the sum of the unsuitability numbers. We do this by changing the signs of all the entries, which gives the right-hand side table above. Since we choose exactly one entry from each row and column, it is clear that if the same integer is added to (or subtracted from) all the entries in a row or column, the optimal choices still remain the same — although the sum of the corresponding entries of course changes. In this example we can therefore add 11 to all the entries of the first row and 13, 14 and 15 to the entries of the second, third and fourth rows, respectively, and finally subtract 3 from all the entries in the last column to obtain the table

$$\begin{array}{ccccc} & J_1 & J_2 & J_3 & J_4 \\ a_1 & 0 & 2 & 10 & 4 \\ a_2 & 5 & 0 & 6 & 4 \\ a_3 & 2 & 12 & 0 & 0 \\ a_4 & 5 & 2 & 0 & 10 \end{array}.$$

Notice that we made all the entries nonnegative. Consequently, if we can find four 0's, no two in the same row or column, they correspond to an optimal assignment. In this example this is possible in a unique way: the unique optimal assignment is to choose a_1 to J_1, a_2 to J_2, a_3 to J_4 and a_4 to J_3. □

We were lucky in the previous example, because we immediately found four suitable zeroes. In general, we need one more trick.

Theorem 1.8 *Let* $\mathbf{A}$ *be an* $m \times n$ *matrix. The maximum number of independent zeros, i.e., zeros no two in the same row or column, is equal to the minimum number of rows and columns required to cover all the zeros in* $\mathbf{A}$.

Proof. Denote by α the maximum number of independent zeros and by β the minimum number of rows or columns required to cover all the zeros. Clearly, $\beta \geq \alpha$, because we can find α independent zeros in $\mathbf{A}$, and any row or column covers at most one of them.

We need to prove that $\alpha \geq \beta$. Assume that some a rows and b columns cover all the zeros and $a + b = \beta$. Because permuting the rows and columns changes neither α nor β, we may assume that the first a rows and the first b columns cover the zeros. Write $\mathbf{A}$ in the form

$$\mathbf{A} = \begin{pmatrix} \mathbf{C}_{a\times b} & \mathbf{D}_{a\times(n-b)} \\ \mathbf{E}_{(m-a)\times b} & \mathbf{F}_{(m-a)\times(n-b)} \end{pmatrix}. \tag{1.9}$$

We know that there are no zeros in $\mathbf{F}$. We show that there are a independent zeros in $\mathbf{D}$. The same argument shows — by symmetry — that there are b independent zeros in $\mathbf{E}$. Together these $a + b$ zeros, which are clearly independent, show that $\alpha \geq a + b = \beta$.

We use Hall's theorem. Define

$$S_i = \{j \mid d_{ij} = 0\} \subseteq \{1, 2, \ldots, n - b\},$$

the set of locations of the zeros in the i-th row of $\mathbf{D} = (d_{ij})$. We claim that the family $S_1, S_2, \ldots, S_a$ has a system of distinct representatives, i.e., we can choose one zero from each row, no two in the same column. Otherwise, Hall's theorem tells us that the union of some k of these sets has less than k elements, which means that the zeros in these rows can all covered by some $m < k$ columns. But then we obtain a covering of all the zeros in $\mathbf{A}$ which contains fewer than $a + b$ rows and columns, a contradiction. □

Consider now the general optimal assignment problem. Assume that n persons $a_1, a_2, \ldots, a_n$ are to be assigned to n jobs $J_1, J_2, \ldots, J_n$ and that we have an $n \times n$ matrix with integer entries representing the suitability of the applicants to the jobs. In the same way as in Example 1.7 we get to a situation where we have an $n \times n$ matrix $\mathbf{A}$ with nonnegative integer entries representing the *unsuitability* of the applicants, and there is already at least one zero in each row and column.

If there are n independent zeros in $\mathbf{A}$, the problem has been solved: each such set of n independent gives an optimal assignment, and there are no others. Assume that the maximum number of independent zeros in $\mathbf{A}$ is $r < n$. By the previous theorem there are nonnegative integers a and b with $a + b = r$ such that all the zeros in $\mathbf{A}$ can be covered by some a rows and b columns. If they are again the first a rows and b columns and $\mathbf{A}$ is as in (1.9), then denote by s the smallest of the numbers in $\mathbf{F}$. We now add s to all these a rows (i.e., to the entries in $\mathbf{C}$ and $\mathbf{D}$) and subtract s from all except the first b columns (i.e., from the entries in $\mathbf{D}$ and $\mathbf{F}$). The net effect is that all the entries in $\mathbf{C}$ have increased by s and all the entries in $\mathbf{F}$ have decreased by s and all the other entries remained unchanged. Consequently, *the sum of all the entries in the matrix* has decreased by the quantity

$$(n-a)(n-b)s - abs = n(n-a-b)s > 0,$$

because $a + b = r < n$. All the entries in the resulting matrix $\mathbf{A}'$ are still nonnegative integers. If $\mathbf{A}'$ still does not contain n independent zeros, we apply the same trick again. Eventually the process has to terminate, because the sum of all the entries in the matrix decreases in each step and can never become negative.

Example 1.10 Assume that in an optimal assignment problem for four persons and four jobs the unsuitability matrix is

$$\begin{array}{cccc} 6 & 8 & 2 & 7 \\ 5 & 8 & 13 & 9 \\ 2 & 7 & 8 & 9 \\ 4 & 11 & 7 & 10 \end{array}.$$

By subtracting the smallest entry from each row and then the smallest entry from each column, we obtain the matrices

$$\begin{array}{cccc} 4 & 6 & 0 & 5 \\ 0 & 3 & 8 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 7 & 3 & 6 \end{array} \qquad \begin{array}{c|ccc} 4 & 3 & 0 & 1 \\ 0 & 0 & 8 & 0 \\ \hline 0 & 2 & 6 & 3 \\ 0 & 4 & 3 & 2 \end{array}.$$

All the zeros are covered by the first two rows and the first column, so there are no four independent zeros. The smallest entry in the lower right-hand block is 2, so we subtract 2 from each element in this block, and add 2 to the two entries in the upper left-hand block to obtain

$$\begin{array}{cccc} 6 & 3 & 0 & 1 \\ 2 & 0 & 8 & 0 \\ 0 & 0 & 4 & 1 \\ 0 & 2 & 1 & 0 \end{array}.$$

Now we can find four independent zeros. In fact we see that there are exactly two different optimal assignments. In both cases the sum of the unsuitability numbers (in the original matrix) is 22 — in the resulting matrix it is of course 0. □

Notice that the same technique can also be used even if there are more jobs than applicants: we can add dummy extra applicants who are equally suited for all the jobs.

Chapter 2

2-designs

2.1 Definition and basic properties of (v, k, λ) designs

The use of combinatorial objects called designs originates from statistical applications. Let us assume that we wish to compare v varieties of coffee. In order to make the testing procedure as fair as possible it is natural to require that 1) each person participating tastes the same number (say k) of varieties so that each person's opinion has the same weight; and 2) each pair of varieties is compared by the same number of persons (say λ) so that each variety gets the same treatment. One possibility would be to let everyone taste all the varieties. But if v is large, this is very impractical, and the comparisons become rather unreliable. So we try to design the experiment so that $k < v$.

Definition 2.1 *Let $S = \{1, 2, \ldots, v\}$. A collection $\mathcal{D}$ of distinct subsets of S is called a (v, k, λ) design if $2 \leq k < v$, $\lambda > 0$, and*

> *1) each set in $\mathcal{D}$ contains exactly k elements,*
>
> *2) each 2-element subset of S is contained in exactly λ of the sets in $\mathcal{D}$.*

The sets of $\mathcal{D}$ are called blocks, and the number of blocks in $\mathcal{D}$ is denoted by b. The set S is called the base set.

Example 2.2 Let $v = 7$ and $S = \{1, 2, 3, 4, 5, 6, 7\}$. The sets $\{1, 2, 4\}$, $\{2, 3, 5\}$, $\{3, 4, 6\}$, $\{4, 5, 7\}$, $\{5, 6, 1\}$, $\{6, 7, 2\}$ and $\{7, 1, 3\}$ form a $(7, 3, 1)$ design as can easily be verified. Notice the very simple structure this design has: all the blocks are "cyclic shifts" of the first block. $\square$

Theorem 2.3 *If $\mathcal{D}$ is a (v, k, λ) design, then each element of the base set occurs in r blocks, where*

$$r(k-1) = \lambda(v-1). \tag{2.4}$$

Moreover,

$$bk = vr. \tag{2.5}$$

Proof. Let $a \in S$ be fixed and assume that a occurs in r_a blocks. We count in two ways the cardinality of the set

$$\{(a, x, B) \mid B \in \mathcal{D}, a, x \in B, a \neq x\}.$$

For each of the $v-1$ possibilities for x ($x \neq a$) there are exactly λ blocks B containing both a and x. The cardinality of the set is therefore $(v-1)\lambda$. On the other hand, for each of the r_a blocks B containing a, the element x can be chosen to be any of the $k-1$ elements in B other than a. Hence $(v-1)\lambda = r_a(k-1)$. This shows that r_a is independent of the choice of a and proves (2.4).

To prove the second claim we count in two ways the cardinality of the set

$$\{(x, B) \mid B \in \mathcal{D}, x \in B\}.$$

For each $x \in S$ the block B can be chosen in r ways; on the other hand, for each of the b blocks B the element $x \in B$ can be chosen in k ways. Hence $vr = bk$. □

Theorem 2.3 shows that the five parameters v, k, λ, b, r are not independent of each other: we can determine b and r from v, k and λ.

A basic question in design theory is to determine for which values of v, k and λ there is a (v, k, λ) design. Certainly, such designs do not exist for all v, k and λ, already by Theorem 2.3.

Example 2.6 There is no $(11, 6, 2)$ design. Otherwise, Theorem 2.3 implies that $r = 4$ and $6b = 44$, a contradiction. □

Theorem 2.7 *If $\mathcal{D}$ is a (v, k, λ) design, then its complement $\overline{\mathcal{D}}$ defined by*

$$\overline{\mathcal{D}} = \{S \setminus B \mid B \in \mathcal{D}\}$$

is a $(v, v-k, b-2r+\lambda)$ design provided that $b - 2r + \lambda > 0$.

Proof. Clearly every block of $\overline{\mathcal{D}}$ has $v - k$ elements. Moreover, a pair (x, y), $x, y \in S$, $x \neq y$, is contained in $S \setminus B$ if and only if B contains neither x nor y. The number of blocks of $\mathcal{D}$ containing neither x nor y is $b - 2r + \lambda$ by the principle of inclusion and exclusion. □

2.2 Resolvable designs

Definition 2.8 *A (v, k, λ) design $\mathcal{D}$ is resolvable, if $\mathcal{D}$ can be partitioned into r collections $\mathcal{D}_i$ each consisting of $b/r = v/k$ of the blocks and every element of S appears in exactly one block in $\mathcal{D}_i$ for all i. The subsets $\mathcal{D}_i$ are called parallel classes.*

Suppose we have a football league of $2n$ teams and each team plays exactly once against every other team. We wish to arrange the league schedule so that all the matches are played during $2n-1$ days, and on each of these days every team plays one match. In other words, we wish to construct a resolvable $(2n, 2, 1)$ design.

Theorem 2.9 *For every positive n there exists a resolvable $(2n, 2, 1)$ design.*

Proof. It is often convenient to use a base set other than $\{1, 2, \ldots, v\}$. We now take $S = \{\infty, 1, 2, \ldots, 2n-1\}$ as the base set. We have to show how to partition the set $\mathcal{D}$ of all 2-element subsets of S into $2n-1$ parallel classes $\mathcal{D}_1, \ldots, \mathcal{D}_{2n-1}$. Define $\{i, \infty\} \in \mathcal{D}_i$, and $\{a, b\} \in \mathcal{D}_i$, if

$$a + b \equiv 2i \pmod{2n-1}$$

for $a, b \in S \setminus \{\infty\}$. Clearly each 2-element subset of S belongs to a unique $\mathcal{D}_i$ (because $\gcd(2, 2n-1) = 1$); and the unique block in $\mathcal{D}_i$ containing a is $\{a, b\}$ where $b \equiv 2i - a \pmod{2n-1}$ if $a \neq i$, and $\{i, \infty\}$ if $a = i$. □

2.3 Incidence matrix of a design

Definition 2.10 *If $\mathcal{D}$ is a (v, k, λ) design, then the binary $b \times v$ matrix $\mathbf{A} = (a_{ij})$, where*

$$a_{ij} = \begin{cases} 1, & \text{if the } i\text{-th block contains } j, \\ 0, & \text{otherwise,} \end{cases}$$

is called an incidence matrix of the design.

Of course, such a matrix is by no means unique, but depends on the order in which we write the blocks. By definition, each row contains k 1's, and according to Theorem 2.3 each column contains r 1's. Condition 2) in Definition 2.1 means that in if we pick any two columns there are exactly λ rows in which there is 1 in both these columns.

Theorem 2.11 *If $\mathbf{A}$ is an incidence matrix of a (v, k, λ) design, then*

$$\mathbf{A}^T\mathbf{A} = (r - \lambda)\mathbf{I} + \lambda\mathbf{J},$$

where $\mathbf{I}$ is the $v \times v$ identity matrix and $\mathbf{J}$ the $v \times v$ matrix in which every entry is 1.

Proof. Clearly $\mathbf{A}^T\mathbf{A}$ is a $v \times v$ matrix whose (i,j) entry is the real inner product of the i-th and j-th columns of $\mathbf{A}$. If $i = j$, this is just the number of 1's in this column, i.e., equal to r. If $i \neq j$, then it is the number of rows in which both the i-th and j-th column have 1, i.e., it equals λ. □

Theorem 2.12 *(Fisher's inequality) If there is a (v, k, λ) design, then $b \geq v$.*

Proof. Let $\mathbf{A}$ be an incidence matrix of a (v, k, λ) design, and consider the determinant of $\mathbf{A}^T\mathbf{A}$, the matrix we calculated in the previous theorem. By subtracting the first row from the others we obtain

$$\det(\mathbf{A}^T\mathbf{A}) = \begin{vmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda & r & \lambda & \dots & \lambda \\ \lambda & \lambda & r & \dots & \lambda \\ \vdots & & & & \vdots \\ \lambda & \lambda & \lambda & \dots & r \end{vmatrix} = \begin{vmatrix} r & \lambda & \lambda & \dots & \lambda \\ \lambda - r & r - \lambda & 0 & \dots & 0 \\ \lambda - r & 0 & r - \lambda & \dots & 0 \\ \vdots & & & & \vdots \\ \lambda - r & 0 & 0 & \dots & r - \lambda \end{vmatrix}.$$

By adding the other columns to the first one we obtain

$$\begin{aligned} \det(\mathbf{A}^T\mathbf{A}) &= \begin{vmatrix} r + (v-1)\lambda & \lambda & \lambda & \dots & \lambda \\ 0 & r - \lambda & 0 & \dots & 0 \\ 0 & 0 & r - \lambda & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & r - \lambda \end{vmatrix} \\ &= (r + (v-1)\lambda)(r - \lambda)^{v-1} \\ &= rk(r - \lambda)^{v-1} \end{aligned}$$

by (2.5), which also implies that $r > \lambda$, because we have assumed $k < v$. Therefore $\det(\mathbf{A}^T\mathbf{A}) \neq 0$.

Assume now on the contrary that $b < v$. Then there are fewer rows than columns in $\mathbf{A}$. We add $v - b$ rows of zeros to $\mathbf{A}$ to obtain a $v \times v$ matrix $\mathbf{A}_1$. Clearly $\mathbf{A}_1^T\mathbf{A}_1 = \mathbf{A}^T\mathbf{A}$, because adding zeros has no effect on the real inner products of the columns. But since $\mathbf{A}_1$ is a square matrix, the product rule for determinants implies that

$$\det(\mathbf{A}^T\mathbf{A}) = \det(\mathbf{A}_1^T\mathbf{A}_1) = \det(\mathbf{A}_1^T)\det(\mathbf{A}_1) = 0,$$

because there is at least one row of zeros in $\mathbf{A}_1$. □

2.4 Symmetric designs

Definition 2.13 *A (v, k, λ) design is symmetric, if $b = v$, i.e., its incidence matrix is a square matrix. For a symmetric design, we denote $n = k - \lambda$. If $n = 1$, the symmetric design is called trivial.*

Clearly, the condition $b = v$ is equivalent to $k = r$ (by (2.5)).

Assume that $\mathcal{D}$ is a symmetric (v, k, λ) design. Because $k < v$ and $\lambda(v - 1) = k(k-1)$ (by (2.4)), we see that $\lambda < k$, i.e., $n = k - \lambda \geq 1$. Moreover, equality holds if and only if $k = v - 1$ and since $b = v$, the design consists of all the $(v-1)$-element subsets of S. Consequently, $k \leq v - 2$ for a nontrivial symmetric design.

Theorem 2.14 *If $\mathbf{A}$ is an incidence matrix of a symmetric (v, k, λ) design, then $\mathbf{A}\mathbf{A}^T = \mathbf{A}^T\mathbf{A}$. In particular, the intersection of any two blocks has cardinality λ.*

Proof. The following clever proof is based on the fact that a matrix commutes with its inverse. Clearly $\mathbf{AJ} = \mathbf{JA} = k\mathbf{J}$ and $\mathbf{A}^T\mathbf{J} = \mathbf{J}\mathbf{A}^T = k\mathbf{J}$, and trivially $\mathbf{J}^2 = v\mathbf{J}$. By Theorem 2.11 (where now $r = k$),

$$(\mathbf{A}^T - \sqrt{\frac{\lambda}{v}}\mathbf{J})(\mathbf{A} + \sqrt{\frac{\lambda}{v}}\mathbf{J}) = \mathbf{A}^T\mathbf{A} + \sqrt{\frac{\lambda}{v}}(\mathbf{A}^T\mathbf{J} - \mathbf{J}\mathbf{A}) - \frac{\lambda}{v}\mathbf{J}^2 = \mathbf{A}^T\mathbf{A} - \lambda\mathbf{J} = (k - \lambda)\mathbf{I}.$$

Consequently $\frac{1}{k-\lambda}(\mathbf{A} + \sqrt{\frac{\lambda}{v}}\mathbf{J})$ is the inverse of the matrix $\mathbf{A}^T - \sqrt{\frac{\lambda}{v}}\mathbf{J}$, and because they commute, we get

$$(k - \lambda)\mathbf{I} = (\mathbf{A} + \sqrt{\frac{\lambda}{v}}\mathbf{J})(\mathbf{A}^T - \sqrt{\frac{\lambda}{v}}\mathbf{J}) = \mathbf{A}^T\mathbf{A} + \sqrt{\frac{\lambda}{v}}(\mathbf{J}\mathbf{A}^T - \mathbf{A}\mathbf{J}) - \frac{\lambda}{v}\mathbf{J}^2 = \mathbf{A}\mathbf{A}^T - \lambda\mathbf{J},$$

i.e., $\mathbf{A}\mathbf{A}^T = (k - \lambda)\mathbf{I} + \lambda\mathbf{J} = \mathbf{A}^T\mathbf{A}$. □

It should be pointed out that the incidence matrix of a symmetric design *need not be symmetric*, cf. Example 2.2. In an incidence matrix $\mathbf{A}$ of a symmetric (v, k, λ) design not just each row but also each column has exactly k 1's, and not just every two columns but also every two rows have λ 1's in common. Therefore also $\mathbf{A}^T$ is an incidence matrix for some (v, k, λ) design. Notice that by Fisher's inequality, the transpose of an incidence matrix design could be an incidence matrix of a design only if $b = v$.

Theorem 2.15 *Assume that $\mathcal{D}$ is a nontrivial symmetric (v, k, λ) design. Then $v - 2k + \lambda > 0$, and the complement $\overline{\mathcal{D}}$ is a symmetric $(\overline{v}, \overline{k}, \overline{\lambda})$ design, where $(\overline{v}, \overline{k}, \overline{\lambda}) = (v, v - k, v - 2k + \lambda)$. Moreover, $\overline{n} = \overline{k} - \overline{\lambda} = k - \lambda = n$.*

Proof. By the discussion preceding the proof $k \leq v-2$. Let B be any block and $\{x,y\} \subseteq S \setminus B$, $x \neq y$. By the proof of Theorem 2.7, $v-2k+\lambda$ is the number of blocks of $\mathcal{D}$ containing neither x nor y, and hence positive. The result now follows from Theorem 2.7. □

Theorem 2.16 *If a nontrivial symmetric (v,k,λ) design exists, then*

$$4n-1 \leq v \leq n^2+n+1.$$

Proof. Using the notations of Theorem 2.15, $\lambda+\overline{\lambda} = v-2n$ and

$$\begin{aligned} \lambda\overline{\lambda} &= \lambda(v-2k+\lambda) = \lambda(v-1)+\lambda-2\lambda k+\lambda^2 = k(k-1)+\lambda-2\lambda k+\lambda^2 \\ &= (k-\lambda)^2-(k-\lambda) = n(n-1). \end{aligned}$$

For all real numbers x and y, $(x+y)^2 \geq 4xy$. Hence for $x=\lambda$ and $y=\overline{\lambda}$, $(v-2n)^2 \geq 4n(n-1) > (2n-2)^2$ because $n \geq 2$, and therefore $(v-2n)^2 \geq (2n-1)^2$. Furthermore $v-2n = \lambda+\overline{\lambda} > 0$, and hence $v-2n \geq 2n-1$, i.e., $v \geq 4n-1$.

Because $\lambda \geq 1$ and $\overline{\lambda} \geq 1$,

$$0 \leq (\lambda-1)(\overline{\lambda}-1) = \lambda\overline{\lambda}-(\lambda+\overline{\lambda})+1 = n(n-1)-(v-2n)+1,$$

i.e., $v \leq n^2+n+1$. □

Consider now the extreme cases.

Theorem 2.17 *If $\mathcal{D}$ is a nontrivial symmetric (v,k,λ) design with $v=n^2+n+1$, then $\mathcal{D}$ or $\overline{\mathcal{D}}$ is a $(n^2+n+1,n+1,1)$ design.*

Proof. From the proof of Theorem 2.16 we see that $v=n^2+n+1$ implies that $\lambda=1$ or $\overline{\lambda}=1$. If $\lambda=1$, then $k=n+\lambda=n+1$, and hence $\mathcal{D}$ is a $(n^2+n+1,n+1,1)$ design. If $\overline{\lambda}=1$, then $\overline{k}=\overline{n}+\overline{\lambda}=\overline{n}+1=n+1$ by Theorem 2.15, and hence $\overline{\mathcal{D}}$ is a $(n^2+n+1,n+1,1)$ design. □

Definition 2.18 *An $(n^2+n+1,n+1,1)$ design (which is automatically symmetric) is called a (finite) projective plane of order n.*

In fact, the finite projective planes are the only symmetric designs with $\lambda=1$. Indeed, if $\lambda=1$, then $k=n+\lambda=n+1$, and $v-1=\lambda(v-1)=k(k-1)=(n+1)n$, i.e., $v=n^2+n+1$.

Theorem 2.19 *If $\mathcal{D}$ is a nontrivial symmetric (v,k,λ) design with $v=4n-1$, then $\mathcal{D}$ or $\overline{\mathcal{D}}$ is a $(4n-1,2n-1,n-1)$ design.*

Proof. From the proof of Theorem 2.16 we see that $\lambda\overline{\lambda} = n(n-1)$ and $\lambda+\overline{\lambda} = v-2n = 2n-1$. Thus λ and $\overline{\lambda}$ are the roots of the quadratic equation $x^2-(2n-1)x+n(n-1) = 0$, i.e., they are n and $n-1$. If $\lambda = n-1$, then $k = n+\lambda = 2n-1$, and $\mathcal{D}$ is a $(4n-1, 2n-1, n-1)$ design. If $\overline{\lambda} = n-1$, then $\overline{k} = \overline{n}+\overline{\lambda} = n+\overline{\lambda} = 2n-1$, and $\overline{\mathcal{D}}$ is a $(4n-1, 2n-1, n-1)$ design.□

Definition 2.20 *An $(4n-1, 2n-1, n-1)$ design (which is automatically symmetric) is called a Hadamard design of order n.*

2.5 Hadamard matrices and designs

Definition 2.21 *An $m \times m$ $(m \geq 2)$ matrix $\mathbf{H}$ whose entries belong to the set $\{-1, 1\}$ is called a Hadamard matrix of order m if $\mathbf{H}^T\mathbf{H} = m\mathbf{I}$.*

Example 2.22 We clearly obtain an infinite family of Hadamard matrices by defining

$$\mathbf{H}_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \mathbf{H}_{2m} = \begin{pmatrix} \mathbf{H}_m & \mathbf{H}_m \\ \mathbf{H}_m & -\mathbf{H}_m \end{pmatrix}$$

for $m = 2, 4, 8, \ldots$.

If we know that all the entries in $\mathbf{H}$ belong to the set $\{-1, 1\}$, then the equation $\mathbf{H}^T\mathbf{H} = m\mathbf{I}$ is equivalent to saying that the columns of $\mathbf{H}$ are orthogonal (i.e., their real inner product is 0).

Let $\mathbf{H}$ be an $m \times m$ matrix. If $\mathbf{H}^T\mathbf{H} = m\mathbf{I}$, then $\mathbf{H}\mathbf{H}^T = m\mathbf{I}$; and conversely. Indeed, in either case we see that $\mathbf{H}$ has an inverse and $\mathbf{H}^{-1} = \frac{1}{m}\mathbf{H}^T$, and the matrix and its inverse commute. For a matrix $\mathbf{H}$ whose entries belong to the set $\{-1, 1\}$, the equation $\mathbf{H}\mathbf{H}^T = m\mathbf{I}$ is equivalent to saying that the rows of $\mathbf{H}$ are orthogonal.

Theorem 2.23 *An $m \times m$ matrix whose elements belong to the set $\{-1, 1\}$ is a Hadamard matrix if and only if its columns (rows) are orthogonal.* □

We can therefore multiply any rows and columns by -1 to obtain other Hadamard matrices. A Hadamard matrix is *normalized* if its first row and column consist entirely of 1's.

Theorem 2.24 *If $\mathbf{H}$ is a Hadamard matrix of order m and its first row (column) consists entirely of 1's, then every other row (column) has $m/2$ 1's and $m/2$ -1's. If $m > 2$, then any two rows (columns) other than the first row (column) have exactly $m/4$ 1's in common.*

Proof. The first statement immediately follows from the fact that the inner product of any row with the first row is 0.

Let R and S be two rows other than the first, and u (resp. v) the number of places where they both have 1's (resp. -1's). Because R has $m/2$ 1's and $m/2$ -1's we get the figure

$$\begin{array}{lcccc}
\text{First row} & 1\,1\ldots1 & 1\;\;1\ldots\;\;1 & 1\;\;1\ldots\;\;1 & 1\;\;1\ldots\;\;1 \\
R & 1\,1\ldots1 & 1\;\;1\ldots\;\;1 & -1\,-1\ldots-1 & -1\,-1\ldots-1 \\
S & 1\,1\ldots1 & -1\,-1\ldots-1 & 1\;\;1\ldots\;\;1 & -1\,-1\ldots-1 \\
 & u & m/2-u & m/2-v & v
\end{array} .$$

Because S has $m/2$ 1's, the third quantity $m/2-v$ has to be equal to $m/2-u$, i.e., $u=v$. The orthogonality of R and S then implies $u-(m/2-u)-(m/2-u)+u=0$, i.e., $u=m/4$.

The claim for columns immediately follows by applying the result for rows to the the Hadamard matrix $\mathbf{H}^T$. □

Corollary 2.25 *If there is a Hadamard matrix of order m, then $m=2$ or m is divisible by 4.* □

It is conjectured that Hadamard matrices exist for *all* orders that are divisible by 4.

Theorem 2.26 *There exists a Hadamard matrix of order $4n$ if and only if there exists a Hadamard design of order n, i.e., a $(4n-1, 2n-1, n-1)$ design.*

Proof. Assume first that there exists a Hadamard matrix of order $4n$, and let $\mathbf{H}$ be a normalized Hadamard matrix of order $4n$. Form a $(4n-1)\times(4n-1)$ matrix $\mathbf{A}$ by deleting the first row and column in $\mathbf{H}$ and changing -1's to 0's. This is an incidence matrix of a $(4n-1, 2n-1, n-1)$ design, because by Theorem 2.24 each row of $\mathbf{A}$ has $2n-1$ 1's and any two columns of $\mathbf{A}$ have exactly $n-1$ 1's in common.

Conversely, assume that there exists a $(4n-1, 2n-1, n-1)$ design, and let $\mathbf{A}$ be its incidence matrix. Form a matrix $\mathbf{H}$ by changing the 0's in $\mathbf{A}$ to -1's and adding a further row and column of 1's. Consider the inner product of the i-th and j-th columns $(1 < i < j)$. The number of positions in which they both have 1 equals $1+\lambda = 1+(n-1) = n$. Since there are $2n$ 1's in each column, there are n positions where the i-th column has 1 and the j-th column has -1 and similarly n positions where the i-th column has -1 and the j-th column has 1. All in all, their inner product is $n-n-n+n=0$. Therefore $\mathbf{H}$ is a Hadamard matrix by Theorem 2.23. □

Let $\mathbf{H}$ be a Hadamard matrix of order m. Take all the rows of the matrices $\mathbf{H}$ and $-\mathbf{H}$, and change all -1's to 0. In this way we obtain a set of $2m$ binary vectors of length m called the *Hadamard code* $\mathcal{B}_m$.

Theorem 2.27 *Every two codewords in $\mathcal{B}_m$ differ in at least $m/2$ coordinates.*

Proof. Take any $\mathbf{a}, \mathbf{b} \in \mathcal{B}_m$, $\mathbf{a} \neq \mathbf{b}$. If $\mathbf{a}$ and $\mathbf{b}$ have been obtained from the i-th rows of $\mathbf{H}$ and $-\mathbf{H}$ respectively, then $\mathbf{a}$ disagrees with $\mathbf{b}$ in all m coordinates. Otherwise for some two different rows $\mathbf{x}$ and $\mathbf{y}$ of $\mathbf{H}$, the word $\mathbf{a}$ is obtained (by changing -1's to 0's) from $\mathbf{x}$ or $-\mathbf{x}$, and $\mathbf{b}$ from $\mathbf{y}$ or $-\mathbf{y}$. In all cases, $\mathbf{a}$ and $\mathbf{b}$ differ in $m/2$ coordinates, because $\mathbf{x}$ and $\mathbf{y}$ are orthogonal, $\mathbf{x}$ and $-\mathbf{y}$ are orthogonal, $-\mathbf{x}$ and $\mathbf{y}$ are orthogonal, and $-\mathbf{x}$ and $-\mathbf{y}$ are orthogonal. □

2.6 Finite projective planes

Recall that a commutative ring $(F, +, \cdot)$ where each nonzero element a has a multiplicative inverse a^{-1} is called a *field*. For instance, if p is a prime, the set of residue classes $\mathbb{Z}_p$ modulo p with respect to usual addition and multiplication is a field. From algebra we assume the following result.

Theorem 2.28 *For every prime power q there exists a finite field $\mathbb{F}_q$ of q elements.*

In the definition of a (v, k, λ) design we have used $S = \{1, 2, \ldots, v\}$ as the base set. The choice of the base set is of course irrelevant, and we can choose any set of v distinct elements just as well.

Denote by V the set of all vectors $\mathbf{x} = (x_0, x_1, x_2)$ of elements in $\mathbb{F}_q$ where x_0, x_1, x_2 are not all zero. We identify the vectors that can obtained from each other by multiplying by an element in $\mathbb{F}_q^* = \mathbb{F}_q \setminus \{0\}$. More formally, we define an equivalence relation $\sim$ in the set V by the condition

$$\mathbf{x} \sim \mathbf{y} \text{ if } \mathbf{x} = \lambda \mathbf{y} \text{ for some } \lambda \in \mathbb{F}_q^*.$$

Clearly, this is an equivalence relation. We denote the equivalence class containing $\mathbf{x}$ by $[\mathbf{x}]$, and the set of all equivalence classes by S.

Example 2.29 Take $q = 3$; then $\mathbb{F}_3 = \mathbb{Z}_3$, and we denote the three elements in $\mathbb{F}_3$ by 0,1 and 2. The set V consists of all the 26 nonzero vectors $(0,0,1)$, $(0,0,2)$, $(0,1,0)$, $(0,1,1)$, $(0,1,2) \ldots, (2,2,2)$. Now $\mathbb{F}_q^* = \{1, 2\}$, so we identify two vectors $\mathbf{x}$ and $\mathbf{y}$ if $\mathbf{x} = \mathbf{y}$ or $\mathbf{x} = 2\mathbf{y}$. Consequently, each vector is identified with exactly one other vector and the set S consists of the 13 equivalence classes

$$
\begin{aligned}
[(0,0,1)] &= [(0,0,2)] = \{(0,0,1),(0,0,2)\}\\
[(0,1,0)] &= [(0,2,0)] = \{(0,1,0),(0,2,0)\}\\
[(0,1,1)] &= [(0,2,2)] = \{(0,1,1),(0,2,2)\}\\
[(0,1,2)] &= [(0,2,1)] = \{(0,1,2),(0,2,1)\}\\
[(1,0,0)] &= [(2,0,0)] = \{(1,0,0),(2,0,0)\}\\
[(1,0,1)] &= [(2,0,2)] = \{(1,0,1),(2,0,2)\}\\
[(1,0,2)] &= [(2,0,1)] = \{(1,0,2),(2,0,1)\}\\
[(1,1,0)] &= [(2,2,0)] = \{(1,1,0),(2,2,0)\}\\
[(1,1,1)] &= [(2,2,2)] = \{(1,1,1),(2,2,2)\}\\
[(1,1,2)] &= [(2,2,1)] = \{(1,1,2),(2,2,1)\}\\
[(1,2,0)] &= [(2,1,0)] = \{(1,2,0),(2,1,0)\}\\
[(1,2,1)] &= [(2,1,2)] = \{(1,2,1),(2,1,2)\}\\
[(1,2,2)] &= [(2,1,1)] = \{(1,2,2),(2,1,1)\}.
\end{aligned}
$$

□

In general, each equivalence class consists of $q-1$ vectors, and therefore the cardinality of S is $(q^3-1)/(q-1) = q^2+q+1$. We use this set S as our base set. In the following construction it is natural to call the elements of S *points*.

Define the blocks — also called *lines* in this context — as follows: the block $B(\alpha)$, where $\alpha = (\alpha_0, \alpha_1, \alpha_2) \in V$ is defined to be the set of all $[\mathbf{x}]$ such that

$$\alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 = 0. \tag{2.30}$$

Notice that if $\mathbf{x}$ satisfies (2.30), so does $\lambda\mathbf{x}$ for all $\lambda \in \mathbb{F}_q^*$. Our design now consists of all the different blocks that can be obtained in this way. Clearly $B(\alpha) = B(\lambda\alpha)$ for every $\lambda \in \mathbb{F}_q^*$.

Because $\alpha \in V$, α has at least one nonzero component; say $\alpha_0 \neq 0$. Therefore (2.30) has exactly q^2-1 solutions $(x_0, x_1, x_2) \in V$: for arbitrary x_1, x_2, both not zero, the equation (2.30) uniquely determines x_0. Since each $[\mathbf{x}]$ consists of $q-1$ vectors of V, there are exactly $(q^2-1)/(q-1) = q+1$ points $[\mathbf{x}]$ satisfying (2.30). In other words: *there are exactly $q+1$ points on each line.*

Finally, assume that $[\mathbf{x}]$ and $[\mathbf{y}]$ are any two given distinct points. How many lines contain both $[\mathbf{x}]$ and $[\mathbf{y}]$? For such a line $B(\alpha)$,

$$
\begin{aligned}
\alpha_0 x_0 + \alpha_1 x_1 + \alpha_2 x_2 &= 0\\
\alpha_0 y_0 + \alpha_1 y_1 + \alpha_2 y_2 &= 0.
\end{aligned}
$$

Without loss of generality $x_0 \neq 0$. We can then replace the second equation by

$$\alpha_1\left(y_1 - \frac{y_0}{x_0}x_1\right) + \alpha_2\left(y_2 - \frac{y_0}{x_0}x_2\right) = 0. \tag{2.31}$$

If

$$y_1 - \frac{y_0}{x_0}x_1 = y_2 - \frac{y_0}{x_0}x_2 = 0$$

then $(y_0, y_1, y_2) = (y_0/x_0)(x_0, x_1, x_2)$ and $[\mathbf{x}] = [\mathbf{y}]$. Therefore at least one of them, say $y_1 - (y_0/x_0)x_1$, is nonzero. Then for arbitrary nonzero α_2, both α_1 and α_0 are uniquely determined by (2.31) and the first equation; and if $(\alpha_0, \alpha_1, \alpha_2)$ is a solution then $(\lambda\alpha_0, \lambda\alpha_1, \lambda\alpha_2)$ for $\lambda \in \mathbb{F}_q^*$ are all the solutions. Consequently, *every two different points* $[\mathbf{x}]$ *and* $[\mathbf{y}]$ *are contained in a unique line.* The argument also shows that $B(\alpha) = B(\beta)$ if and only if $\alpha \sim \beta$.

Theorem 2.32 *For every prime power q there exists a projective plane of order q, i.e., a $(q^2 + q + 1, q + 1, 1)$ design.* □

The number of blocks is $q^2 + q + 1$: this is clear by the construction, and also follows from (2.4) and (2.5).

Example 2.33 We use the previous example and construct a projective plane of order 3, i.e., a $(13, 4, 1)$ design. From 2.30 we get the 13 lines

line	defining equation	points
$B(1,0,0)$	$x_0 = 0$	$[(0,0,1)], [(0,1,0)], [(0,1,1)], [(0,1,2)]$
$B(0,1,0)$	$x_1 = 0$	$[(0,0,1)], [(1,0,0)], [(1,0,1)], [(1,0,2)]$
$B(0,0,1)$	$x_2 = 0$	$[(0,1,0)], [(1,0,0)], [(1,1,0)], [(1,2,0)]$
$B(1,1,0)$	$x_0 + x_1 = 0$	$[(0,0,1)], [(1,2,0)], [(1,2,1)], [(1,2,2)]$
$B(0,1,1)$	$x_1 + x_2 = 0$	$[(1,0,0)], [(0,1,2)], [(1,1,2)], [(1,2,1)]$
$B(1,0,1)$	$x_0 + x_2 = 0$	$[(0,1,0)], [(1,0,2)], [(1,1,2)], [(1,2,2)]$
$B(1,2,0)$	$x_0 + 2x_1 = 0$	$[(0,0,1)], [(1,1,0)], [(1,1,1)], [(1,1,2)]$
$B(0,1,2)$	$x_1 + 2x_2 = 0$	$[(1,0,0)], [(0,1,1)], [(1,1,1)], [(1,2,2)]$
$B(1,0,2)$	$x_0 + 2x_2 = 0$	$[(0,1,0)], [(1,0,1)], [(1,1,1)], [(1,2,1)]$
$B(1,1,1)$	$x_0 + x_1 + x_2 = 0$	$[(1,1,1)], [(1,0,2)], [(1,2,0)], [(0,1,2)]$
$B(1,2,1)$	$x_0 + 2x_1 + x_2 = 0$	$[(0,1,1)], [(1,1,0)], [(1,0,2)], [(1,2,1)]$
$B(1,1,2)$	$x_0 + x_1 + 2x_2 = 0$	$[(0,1,1)], [(1,0,1)], [(1,2,0)], [(1,1,2)]$
$B(1,2,2)$	$x_0 + 2x_1 + 2x_2 = 0$	$[(0,1,2)], [(1,0,1)], [(1,1,0)], [(1,2,2)]$.

By the previous theorem we know that if we pick any two of the 13 points

$$[(0,0,1)], [(0,1,0)], [(0,1,1)], [(0,1,2)], [(1,0,0)], [(1,0,1)], [(1,0,2)],$$

$$[(1,1,0)], [(1,1,1)], [(1,1,2)], [(1,2,0)], [(1,2,1)], [(1,2,2)],$$

they are contained in a unique line. If we like, we can of course rename the points to $1, 2, \ldots, 13$, in which case our design gets a more familiar look $\{1,2,3,4\}, \{1,5,6,7\}, \ldots, \{4,6,8,13\}$. □

We can more generally consider vectors $\mathbf{x} = (x_0, x_1, \ldots, x_m)$ and define $\mathbf{x} \sim \mathbf{y}$ if $\mathbf{x} = \lambda \mathbf{y}$ for some $\mathbb{F}_q^*$. Let S again denote the set of $(q^{m+1}-1)/(q-1)$ equivalence classes $[\mathbf{x}]$. For every $\alpha \in \mathbb{F}_q^*$, let $B(\alpha)$ be the set of all $[\mathbf{x}]$ such that

$$\alpha_0 x_0 + \alpha_1 x_1 + \ldots + \alpha_m x_m = 0.$$

Exactly the same argument as before shows that each block contains $(q^m - 1)/(q-1)$ points $[\mathbf{x}]$, and every two distinct points $[\mathbf{x}]$ and $[\mathbf{y}]$ are contained in exactly $(q^{m-1}-1)/(q-1)$ lines. The resulting design is denoted by $PG(m,q)$. We therefore have the following generalization of Theorem 2.32.

Theorem 2.34 *For every prime power q there exists a*

$$(\frac{q^{m+1}-1}{q-1}, \frac{q^m-1}{q-1}, \frac{q^{m-1}-1}{q-1})$$

design. □

2.7 Lagrange's theorem

Theorem 2.35 *(Lagrange) Every positive integer can be written as a sum of four squares of integers.*

It is easy to verify that

$$(x_1^2 + x_2^2 + x_3^2 + x_4^2)(y_1^2 + y_2^2 + y_3^2 + y_4^2) = z_1^2 + z_2^2 + z_3^2 + z_4^2,$$

where

$$\begin{aligned} z_1 &= x_1y_1 + x_2y_2 + x_3y_3 + x_4y_4, \\ z_2 &= x_2y_1 - x_1y_2 + x_4y_3 - x_3y_4, \\ z_3 &= x_3y_1 - x_4y_2 - x_1y_3 + x_2y_4, \\ z_4 &= x_4y_1 + x_3y_2 - x_2y_3 - x_1y_4. \end{aligned}$$

Indeed, for arbitrary complex numbers a, b, c, d,

$$\begin{aligned} &|a\overline{c} + b\overline{d}|^2 + |bc - ad|^2 \\ &= (a\overline{c} + b\overline{d})(\overline{a}c + \overline{b}d) + (bc - ad)(\overline{bc} - \overline{ad}) \\ &= a\overline{a}c\overline{c} + b\overline{b}d\overline{d} + b\overline{b}c\overline{c} + a\overline{a}d\overline{d} \\ &= |a|^2|c|^2 + |b|^2|d|^2 + |b|^2|c|^2 + |a|^2|d|^2 \\ &= (|a|^2 + |b|^2)(|c|^2 + |d|^2), \end{aligned}$$

and our the claim now immediately follows by substituting $a = x_1 + x_2 i$ $b = x_3 + x_4 i$, $c = y_1 + y_2 i$, $d = y_3 + y_4 i$, because then $a\overline{c} + b\overline{d} = z_1 + z_2 i$ and $bc - ad = z_3 + z_4 i$.

This implies that if m and n can be expressed as sums of four squares, so can mn. Because $2 = 1^2+1^2+0^2+0^2$ it suffices to prove Lagrange's theorem for an arbitrary odd prime p. We first require a lemma.

Lemma 2.36 *If p is an odd prime, then there exist integers x, y and m such that $1+x^2+y^2 = mp$ and $0 < m < p$.*

Proof. All the elements in the set $A = \{x^2 \mid 0 \le x \le (p-1)/2\}$ belong to different residue classes modulo p. Indeed, for two different elements a^2 and b^2 the difference $a^2 - b^2 = (a-b)(a+b)$ is divisible by p only if $a-b$ or $a+b$ is divisible by p, which is clearly impossible unless $a = b$. Similarly, all the elements of the set $B = \{-1-y^2 \mid 0 \le y \le (p-1)/2\}$ belong to different residue classes modulo p. Both sets have $(p+1)/2$ elements, and therefore we can find some $x^2 \in A$ and $-1-y^2 \in B$ that belong to the same residue class. Hence there is an integer m such that

$$1 + x^2 + y^2 = x^2 - (-1-y^2) = mp,$$

and $0 < 1+x^2+y^2 < 1+2(p/2)^2 < p^2$, so $0 < m < p$. □

Proof of Theorem 2.35. Let p be an odd prime. By Lemma 2.36 there exist integers x_1, x_2, x_3, x_4 and a positive integer $m < p$ such that

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

Let m_0 the smallest of such integers m. By Lemma 2.36, $m_0 < p$. We need to prove that $m_0 = 1$. Suppose $m_0 > 1$.

If m_0 is even, then $x_1^2+x_2^2+x_3^2+x_4^2$ is even and the integers x_1, x_2, x_3, x_4 are all even, all odd, or exactly two of them — say x_1 and x_2 — are even; anyway x_1+x_2, x_1-x_2, x_3+x_4 and x_3-x_4 are all even and

$$\frac{m_0}{2}p = \Big(\frac{x_1+x_2}{2}\Big)^2 + \Big(\frac{x_1-x_2}{2}\Big)^2 + \Big(\frac{x_3+x_4}{2}\Big)^2 + \Big(\frac{x_3-x_4}{2}\Big)^2.$$

Therefore $m_0/2$ can be represented as a sum of four squares, a contradiction. Hence m_0 is odd, and in particular $m_0 \ge 3$.

The integers x_1, x_2, x_3, x_4 cannot all be divisible by m_0; otherwise m_0^2 would divide $m_0 p = x_1^2+x_2^2+x_3^2+x_4^2$ and hence m_0 would divide p although $1 < m_0 < p$. Choose integers b_1, b_2, b_3, b_4 so that

$$y_i = x_i - b_i m_0 \text{ and } |y_i| < m_0/2$$

for $i = 1,2,3,4$. Then at least one of the integers y_1, y_2, y_3, y_4 is nonzero, and

$$0 < y_1^2+y_2^2+y_3^2+y_4^2 < 4(m_0/2)^2 = m_0^2.$$

Since $x_1^2+x_2^2+x_3^2+x_4^2$ and all the differences $y_i^2-x_i^2=b_i^2m_0^2-2x_ib_im_0$ are divisible by m_0, so is $y_1^2+y_2^2+y_3^2+y_4^2$. Hence there exists an integer m_1 such that

$$\begin{aligned} x_1^2+x_2^2+x_3^2+x_4^2 &= m_0p \\ y_1^2+y_2^2+y_3^2+y_4^2 &= m_0m_1, \end{aligned}$$

where $0<m_1<m_0$. Then

$$m_0^2m_1p = z_1^2+z_2^2+z_3^2+z_4^2,$$

where z_1, z_2, z_3, z_4 are as at the beginning of the section. Moreover,

$$z_1=\sum_{i=1}^4 x_iy_i=\sum_{i=1}^4 x_i(x_i-b_im_0)\equiv\sum_{i=1}^4 x_i^2\equiv 0 \pmod{m_0},$$

and in the same way z_2, z_3 and z_4 are divisible by m_0. Consequently, there exist integers t_i such that $z_i=m_0t_i$ $(i=1,2,3,4)$, and we obtain

$$m_1p=t_1^2+t_2^2+t_3^2+t_4^2.$$

But here $0<m_1<m_0<p$, which contradicts the minimality of m_0. This shows that the assumption $m_0>1$ must have been false. □

2.8 The theorem of Bruck, Ryser and Chowla

Theorem 2.37 *If a* **symmetric** (v,k,λ) *design exists and* v *is even, then* $n=k-\lambda$ *is a square.*

Proof. By the proof of Fisher's inequality, the incidence matrix $\mathbf{A}$ of such a design satisfies $\det(\mathbf{A})^2=\det(\mathbf{A}^T)\det(\mathbf{A})=k^2(k-\lambda)^{v-1}$. Here $\det(\mathbf{A})$, k and $k-\lambda$ are all positive. By writing them as products of primes, we see that $k-\lambda$ has to be a square. □

Example 2.38 We show that no $(22,7,2)$ design exists. If such a design does exist, it is symmetric, because $2(22-1)=7(7-1)$. But $k-\lambda=7-2=5$ is not a square, so the design cannot exist by the previous theorem. □

Lemma 2.39 *If* $\mathbf{R}$ *is any* $m\times m$ *matrix over* $\mathbb{Q}$, *then there exists a matrix*

$$\mathbf{E}=\begin{pmatrix} \epsilon_1 & & & \bigcirc \\ & \epsilon_2 & & \\ & & \ddots & \\ \bigcirc & & & \epsilon_m \end{pmatrix}$$

where all $\epsilon_i \in \{-1, 1\}$, such that $\mathbf{R} - \mathbf{E}$ is invertible. □

The case v odd is more difficult.

Theorem 2.40 *(Bruck, Ryser and Chowla) If there exists a* **symmetric** *(v, k, λ) design and v is odd, then the equation*

$$z^2 = (k - \lambda)x^2 + (-1)^{(v-1)/2}\lambda y^2$$

has a nontrivial integer solution $(x, y, z) \neq (0, 0, 0)$.

Proof. By Lagrange's theorem there are integers a, b, c, d such that $n = k - \lambda = a^2 + b^2 + c^2 + d^2$ and then $\mathbf{H}\mathbf{H}^T = \mathbf{H}^T\mathbf{H} = n\mathbf{I}_4$, where

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

Assume that $\mathbf{A}$ is an incidence matrix for a (v, k, λ) design. By Theorem 2.11, $\mathbf{A}^T\mathbf{A} = n\mathbf{I}_v + \lambda\mathbf{J}_v$.

Case $v \equiv 3 \pmod 4$: Denote

$$\mathbf{B} = \begin{pmatrix} & & & 0 \\ & \mathbf{A} & & \vdots \\ & & & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}, \qquad \mathbf{K} = \begin{pmatrix} \mathbf{H} & & & & \\ & \mathbf{H} & & \bigcirc & \\ & & \mathbf{H} & & \\ & \bigcirc & & \ddots & \\ & & & & \mathbf{H} \end{pmatrix},$$

where both are $(v + 1) \times (v + 1)$ matrices. Then

$$\mathbf{B}^T\mathbf{B} = \begin{pmatrix} k & \lambda & \dots & \lambda & 0 \\ \lambda & k & \dots & \lambda & 0 \\ \vdots & & & \vdots & \\ \lambda & \lambda & \dots & k & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix}, \qquad \mathbf{K}^T\mathbf{K} = n\mathbf{I}_{v+1}.$$

Moreover, from the proof of Fisher's inequality, we know that $\mathbf{B}$ is invertible. Denote $\mathbf{P} = \mathbf{B}^{-1}\mathbf{K}$. Clearly, all the entries in $\mathbf{P}$ are rational numbers. If $\mathbf{x} = (x_1, x_2, \dots, x_{v+1})^T \in \mathbb{Q}^{v+1}$ and $\mathbf{y} = (y_1, y_2, \dots, y_{v+1})^T \in \mathbb{Q}^{v+1}$ satisfy the equation $\mathbf{x} = \mathbf{P}\mathbf{y}$, then $\mathbf{B}\mathbf{x} = \mathbf{K}\mathbf{y}$ and

$$n(y_1^2 + y_2^2 + \ldots + y_{v+1}^2) = \mathbf{y}^T\mathbf{K}^T\mathbf{K}\mathbf{y} = \mathbf{x}^T\mathbf{B}^T\mathbf{B}\mathbf{x} \tag{2.41}$$

$$\begin{aligned} &= k(x_1^2 + x_2^2 + \ldots + x_v^2) + x_{v+1}^2 + \lambda \sum_{i,j \le v, i \ne j} x_i x_j \\ &= \lambda(x_1 + x_2 + \ldots + x_v)^2 + x_{v+1}^2 + n(x_1^2 + x_2^2 + \ldots + x_v^2). \end{aligned} \tag{2.42}$$

We now claim that the system $\mathbf{x} = \mathbf{P}\mathbf{y}$ in $2(v+1)$ unknowns x_i and y_j has a solution in $\mathbb{Q}$ such that $y_{v+1} = 1$ and $x_i^2 = y_i^2$ for all $i \le v$. Substitute $y_{v+1} = 1$ in the system $\mathbf{x} = \mathbf{P}\mathbf{y}$. We can omit the last equation: it is the only one involving x_{v+1}. The first v equations form the system

$$\begin{pmatrix} x_1 \\ \vdots \\ x_v \end{pmatrix} = \mathbf{R} \begin{pmatrix} y_1 \\ \vdots \\ y_v \end{pmatrix} + \begin{pmatrix} p_{1,v+1} \\ \vdots \\ p_{v,v+1} \end{pmatrix},$$

where $\mathbf{R}$ is the matrix formed by the first v rows and columns of $\mathbf{P}$. But when we now choose ϵ_i for $i = 1, 2, \ldots, v$ as in Lemma 2.39, and choose $x_i := \epsilon_i y_i$, we obtain a system of linear equations with invertible coefficient matrix $\mathbf{R}-\mathbf{E}$, and we find the required solution.

Substituting this solution to (2.41) and (2.42) we obtain the equation

$$n = ny_{v+1}^2 = \lambda(x_1 + \ldots + x_v)^2 + x_{v+1}^2,$$

where $x_1, \ldots, x_{v+1} \in \mathbb{Q}$. Hence there exist integers $x \ne 0$, y and z such that $n = \lambda(y/x)^2 + (z/x)^2$, and therefore the equation $z^2 = nx^2 - \lambda y^2$ has a nontrivial integer solution, completing the proof in the case $v \equiv 3 \pmod 4$.

Case $v \equiv 1 \pmod 4$: We now use

$$\mathbf{K} = \begin{pmatrix} \mathbf{H} & & & & 0 \\ & \mathbf{H} & & \bigcirc & 0 \\ & & \ddots & & \vdots \\ & \bigcirc & & \mathbf{H} & 0 \\ 0 & & \ldots & 0 & 1 \end{pmatrix}_{v \times v}, \mathbf{K}^T\mathbf{K} = \begin{pmatrix} n & & & & \\ & n & & \bigcirc & \\ & & \ddots & & \\ & \bigcirc & & n & \\ & & & & 1 \end{pmatrix}.$$

Let $\mathbf{x} = (x_1, \ldots, x_v)^T$ and $\mathbf{y} = (y_1, \ldots, y_v)^T$. If $\mathbf{x} = \mathbf{A}^{-1}\mathbf{K}\mathbf{y}$, then

$$n(y_1^2 + \ldots + y_{v-1}^2) + y_v^2 = n(x_1^2 + \ldots + x_v^2) + \lambda(x_1 + \ldots + x_v)^2.$$

As in the case $v \equiv 3 \pmod 4$, we find $x_1, \ldots, x_v \in \mathbb{Q}$ such that $1 = nx_v^2 + \lambda(x_1 + \ldots + x_v)^2$. Hence there exist integers x, y, z, not all zero, such that $z^2 = nx^2 + \lambda y^2$. □

Example 2.43 We show that there is no $(29, 8, 2)$ design. If such a design exists, it is symmetric, because $2(29-1) = 8(8-1)$, and by the theorem of Bruck, Ryser and Chowla, the equation $z^2 = 6x^2 + 2y^2$ has a nontrivial integer solution $(x, y, z) \neq (0, 0, 0)$. Assume without loss of generality that the g.c.d. of x, y and z is 1. Consider the equation modulo 3. We see that $z^2 - 2y^2$ has to be divisible by 3. By the following table $z^2 - 2y^2 \equiv 0 \pmod 3$ if and only if both $y \equiv 0 \pmod 3$ and $z \equiv 0 \pmod 3$.

		$z \equiv$	
$z^2 - 2y^2 \equiv$		0	± 1
$y \equiv$	0	0	1
	± 1	1	2

But if both y and z are divisible by 3, then $6x^2 = z^2 - 2y^2$ is divisible by 9 and hence x is divisible by 3. This is a contradiction, because we assumed that the g.c.d. of x, y and z is 1. Consequently there cannot be any $(29, 8, 2)$ design. □

If there exists a projective plane of order n, then the theorem of Bruck, Ryser and Chowla implies that the equation

$$z^2 = nx^2 + (-1)^{(v-1)/2} y^2$$

has a nontrivial integer solution. If $n \equiv 0, 3 \pmod 4$, then $(v-1)/2 = n(n+1)/2$ is even, and the equation gets the form

$$z^2 = nx^2 + y^2.$$

This equation clearly has nontrivial solutions $x = 0$, $y = z$, and therefore the theorem of Bruck, Ryser and Chowla *gives no information whether or not such a finite projective plane exists.* If $n \equiv 1, 2 \pmod 4$, the equation takes the form

$$y^2 + z^2 = nx^2,$$

and we obtain the following highly nontrivial result.

Theorem 2.44 *If there exists a projective plane of order $n \equiv 1, 2 \pmod 4$, and $n = ab^2$ where a is square-free (i.e., not divisible by the square of any prime), then a has no prime factor $p \equiv 3 \pmod 4$.*

Proof. By the previous discussion we know that the equation

$$y^2 + z^2 = nx^2 = aw^2,$$

where $w = bx$, has a nontrivial integer solution $(x, y, w) \neq (0, 0, 0)$. We may assume that the g.c.d. of y, z and w is 1. Let p be an odd prime factor of a. If $p \mid y$, then $p \mid aw^2 - y^2 = z^2$ and $p \mid z$; consequently $p^2 \mid y^2 + z^2 = aw^2$, but p does not divide w (because the g.c.d. of y, z and w is 1) and a is square-free, a contradiction. Hence p does not divide y, and there exists an integer s such that $sy \equiv 1 \pmod p$. The congruence $y^2 + z^2 \equiv 0 \pmod p$ therefore implies that $(sz)^2 \equiv -1 \pmod p$. We have shown that there is an integer h — obviously not divisible by p — such that $h^2 \equiv -1 \pmod p$. Raising both sides to the power $(p-1)/2$ we get $h^{p-1} \equiv (-1)^{(p-1)/2} \pmod p$. By Fermat's little theorem $h^{p-1} \equiv 1 \pmod p$ and therefore $(-1)^{(p-1)/2} = 1$, i.e., $p \equiv 1 \pmod 4$. □

Example 2.45 There is no projective plane of order 6, because $6 \equiv 2 \pmod 4$ and 6 is square-free, and divisible by 3.

By computer it has been shown that there is no projective plane of order 10. Nothing else is known. It is conjectured that the order of a finite projective plane is a prime power.

2.9 Mutually orthogonal Latin squares

Definition 2.46 *Two Latin squares $A = (a_{ij})$ and $B = (b_{ij})$ of order n are orthogonal if for every pair $(a, b) \in \{(1, 1), (1, 2), \ldots, (n, n)\}$ there exist unique indices i and j such that $(a_{ij}, b_{ij}) = (a, b)$. A set of Latin squares is mutually orthogonal if any two of them are orthogonal.*

Example 2.47 The two Latin squares

$$\begin{array}{cccc} 1 & 2 & 3 & 4 \\ 3 & 4 & 1 & 2 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \end{array} \qquad \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 4 & 3 & 2 & 1 \\ 2 & 1 & 4 & 3 \\ 3 & 4 & 1 & 2 \end{array}$$

are orthogonal. □

Example 2.48 We have already remarked that when defining a Latin square we can use any n symbols as the entries. The same of course applies to the row and column labels. Use now the elements of $\mathbb{Z}_n$ as the n symbols as well as row and column labels. In this case it is natural to view a Latin square as a function $L(x, y)$ from $\mathbb{Z}_n \times \mathbb{Z}_n$ to $\mathbb{Z}_n$. Define

$$L_1(x, y) = x + y, \qquad L_2(x, y) = x - y.$$

These are Latin squares, which for odd n are orthogonal. □

Theorem 2.49 *There are at most $n-1$ mutually orthogonal Latin squares of order n.*

Proof. Assume that we have k mutually orthogonal Latin squares. Let $(a_1, a_2, \ldots, a_n)$ be a permutation of the numbers 1, 2, ..., n. If we apply this permutation to one of the squares, i.e., replace i everywhere with a_i for all $i = 1, 2, \ldots, n$, we again get a Latin square which is orthogonal to all the others. We can therefore assume that the first row in each of the k Latin squares is $(1, 2, \ldots, n)$. Consider the k entries in the position $(2, 1)$. None of them is 1 because 1 already appears in the first column in each square. But no two of these entries can be the same: if s appears twice in the position $(2, 1)$ then the corresponding squares are not orthogonal, because s is also the s-th first row entry in both of them. □

It turns out that the maximum number $n-1$ of orthogonal Latin squares of order n is attained if and only if certain designs exist.

Theorem 2.50 *An affine plane of order n, i.e., an $(n^2, n, 1)$ design, is resolvable.*

Proof. By (2.4) and (2.5), $r = n+1$ and $b = n^2 + n$.

We first show that given any block $B = \{a_1, a_2, \ldots, a_n\}$ and any $x \notin B$, there is a *unique* block which does not intersect B and contains x. Indeed, for every $i = 1, 2, \ldots, n$, there is a unique block B_i containing both x and a_i, and clearly $B_i \neq B_j$ whenever $i \neq j$; otherwise a_i and a_j would be contained in both B and B_i. The remaining of the $r = n+1$ blocks containing x therefore does not intersect B.

In particular, any two blocks not intersecting B are disjoint; otherwise any point x in their intersection would be contained in two blocks not intersecting B. Hence the $n^2 - n$ points not in B form $n-1$ pairwise non-intersecting blocks; and these are the only blocks not intersecting B.

Consequently, we obtain the $n+1$ parallel classes, each consisting of n blocks, by defining that two different blocks are in the same parallel class if they are disjoint. □

Theorem 2.51 *A projective plane of order $n \geq 2$ exists if and only if there is an affine plane of order n.*

Proof. Assume that $\mathcal{D}$ is an affine plane of order n, and $\mathcal{D}_i$, $i = 1, 2, \ldots, n+1$, are its parallel classes. Extend the base set by $n+1$ new elements ∞_1, ∞_2, ..., ∞_{n+1}, and for every i, add the element ∞_i to every block in the i-th parallel class. The new blocks together with the block $\{\infty_1, \infty_2, \ldots, \infty_n\}$ clearly form a projective plane of order n.

Conversely, if $\mathcal{D}$ is a projective plane of order n and $B \in \mathcal{D}$ a given block, we obtain an affine plane of order n by deleting the block B and by deleting from every other block the one element in common with B (the fact that the intersection always consists of exactly one element follows from Theorem 2.14). □

Theorem 2.52 *There exists an affine plane of order n (or equivalently, a projective plane of order n) if and only if there are $n-1$ mutually orthogonal Latin squares of order n.*

Proof. Assume that we have $n-1$ mutually orthogonal Latin squares L_1, $L_2, \ldots, L_{n-1}$ of order n. Form the $(n+1) \times n^2$ array

$$\begin{array}{llll}
1\,1\,1 \ldots\ 1 & 2\,2\,2 \ldots\ 2 & \ldots & n\,n\,n \ldots\ n \\
1\,2\,3 \ldots\ n & 1\,2\,3 \ldots\ n & \ldots & 1\,2\,3 \ldots\ n \\
\text{row 1 in } L_1 & \text{row 2 in } L_1 & \ldots & \text{row } n \text{ in } L_1 \\
\text{row 1 in } L_2 & \text{row 2 in } L_2 & \ldots & \text{row } n \text{ in } L_2 \\
\vdots & & & \\
\text{row 1 in } L_{n-1} & \text{row 2 in } L_{n-1} & \ldots & \text{row } n \text{ in } L_{n-1}
\end{array} .$$

This array has the following orthogonality property: in every two rows all the n^2 vertical pairs $\binom{1}{1}$, $\binom{1}{2}$, ..., $\binom{n}{n}$ appear exactly once: if we compare the i-th and j-th rows, and $i \leq 2 < j$, this is so because the j-th row comes from a Latin square; if $i, j \geq 3$, this follows from the orthogonality of the Latin squares.

Label the columns of the array by $1, 2, \ldots, n^2$. Each of the $n+1$ rows of the array gives us n blocks: for every $i = 1, 2, \ldots, n$ take as a block the set of labels of the columns where the row has i. These $n^2 + n$ blocks form an $(n^2, n, 1)$ design. It remains to prove that $\lambda = 1$. By the orthogonality property of the array, any 2-element subset of the set $\{1, 2, \ldots, n^2\}$ cannot be contained in more than one block. But together the blocks contain $(n^2+n)\binom{n}{2} = \binom{n^2}{2}$ 2-element subsets, i.e., all of them exactly once.

Conversely, given an $(n^2, n, 1)$ design $\mathcal{D}$ with parallel classes $\mathcal{D}_1$, $\mathcal{D}_2$, ..., $\mathcal{D}_{n+1}$, we can (by relabelling the elements of the base set) assume that the first two parallel classes are represented by the first two rows of the array above and write the other $n-1$ parallel classes as the last $n-1$ rows. Because the rows originate from a design with $\lambda = 1$, the resulting array has the orthogonality property described above. But then interpreting the last $n-1$ rows as $n \times n$ squares gives us $n-1$ mutually orthogonal Latin squares. □

In particular, Theorems 2.32 and 2.52 imply that if q is a prime power there exists a set of $q-1$ mutually orthogonal Latin squares of order q.

2.10 Difference sets

Definition 2.53 *A k-element subset $D = \{d_1, d_2, \ldots, d_k\} \subseteq \mathbb{Z}_v$ is called a cyclic (v, k, λ) difference set if $2 \leq k < v$, $\lambda > 0$, and every nonzero $d \in \mathbb{Z}_v$ can be expressed in the form $d = d_i - d_j$ for exactly λ pairs (i, j), $i, j \in \{1, 2, \ldots, k\}$.*

Since the number of pairs (i, j) with $i \neq j$ equals $k(k-1)$ and these give each of the $v - 1$ nonzero elements λ times as a difference, we know that for a cyclic (v, k, λ) difference set

$$\lambda(v-1) = k(k-1). \tag{2.54}$$

If D is a difference set, we call the set $a + D = \{a + d_1, a + d_2, \ldots, a + d_k\}$ a *translate* of D. Notice that our assumption $k < v$ together with (2.54) implies that all the translates of a cyclic difference set are different. Indeed, if $a + D = D$ for some $a \neq 0$, then a can be expressed as a difference in k ways; but $\lambda < k$ by (2.54) and our assumption $k < v$.

Theorem 2.55 *If D is a cyclic (v, k, λ) difference set then the translates D, $1 + D$, $\ldots$, $(v-1) + D$ are the blocks of a symmetric (v, k, λ) design.*

Proof. By the previous discussion we obtain v different k-element blocks. Furthermore, $a, b \in x + D$ $(a \neq b)$ if and only if $a - x = d_i$ and $b - x = d_j$ for some $i \neq j$, i.e., $(a - x, b - x)$ is one of the λ pairs (d_i, d_j) such that $d_i - d_j = a - b$. □

Example 2.56 Let $D = \{1, 2, 4, 5, 6, 10\} \subseteq \mathbb{Z}_{11}$. From the table

		d_j					
	$d_i - d_j$	1	2	4	5	6	10
	1	0	10	8	7	6	2
	2	1	0	9	8	7	3
d_i	4	3	2	0	10	9	5
	5	4	3	1	0	10	6
	6	5	4	2	1	0	7
	10	9	8	6	5	4	0

we see that every nonzero element of $\mathbb{Z}_{11}$ can be expressed as a difference $d_i - d_j$ for exactly three pairs (i, j). Hence D is a cyclic $(11, 6, 3)$ difference set and by Theorem 2.55 the blocks $\{1, 2, 4, 5, 6, 10\}$, $\{2, 3, 5, 6, 7, 11\}$, $\ldots$, $\{1, 3, 4, 5, 9, 11\}$ form a symmetric $(11, 6, 3)$ design. □

Example 2.57 The set $D = \{1, 2, 4\} \subseteq \mathbb{Z}_7$ forms a cyclic $(7, 3, 1)$ difference set and gives us the design of Example 2.2. $\square$

Definition 2.58 *A k-element subset $D = \{d_1, d_2, \ldots, d_k\}$ of an additive abelian group G is called a (v, k, λ) difference set in G if $2 \leq k < v$, $\lambda > 0$, and every nonzero element g of G has exactly λ representations as $g = d_i - d_j$.*

Any difference set in an additive abelian group gives a symmetric design: we take as blocks all the translates $g + D$, $g \in G$.

Theorem 2.59 *If q is a prime power and $q \equiv 3 \pmod 4$, then the nonzero squares in $\mathbb{F}_q$ form a $(q, (q-1)/2, (q-3)/4)$ difference set.*

Proof. Exactly half of the nonzero elements in $\mathbb{F}_q$ are squares. Indeed, the nonzero squares in $\mathbb{F}_q$ are the elements α^2 for $\alpha \in \mathbb{F}_q \setminus \{0\}$, but for every $\alpha \in \mathbb{F}_q \setminus \{0\}$ the equation $x^2 = \alpha^2$ has the two different solutions $x = \pm\alpha$. Denote by $\mathcal{N}$ (resp. $\mathcal{Q}$) the set of nonsquares (resp. nonzero squares) in $\mathbb{F}_q$.

If $q \equiv 3 \pmod 4$, then $-1 \in \mathcal{N}$. Otherwise, $-1 = \alpha^2$ for some $\alpha \in \mathbb{F}_q$, and $-1 = (-1)^{(q-1)/2} = \alpha^{q-1} = 1$ using a result from elementary group theory. Consequently, $\mathcal{N} = -\mathcal{Q} = \{-\gamma \mid \gamma \in \mathcal{Q}\}$.

For any $\gamma \in \mathcal{Q}$, the pair $(x, y) \in \mathcal{Q} \times \mathcal{Q}$ satisfies the equation $x - y = 1$ if and only if the pair $(\gamma x, \gamma y) \in \mathcal{Q} \times \mathcal{Q}$ satisfies the equation $\gamma x - \gamma y = \gamma$, or equivalently if and only if the pair $(\gamma y, \gamma x) \in \mathcal{Q} \times \mathcal{Q}$ satisfies the equation $\gamma y - \gamma x = -\gamma$. This shows that all nonzero squares $\gamma \in \mathcal{Q}$ and all nonsquares $-\gamma \in \mathcal{N}$ have the same number of representations as a difference of two nonzero squares. $\square$

Corollary 2.60 *If $4n - 1$ is a prime power, then there exists a $(4n-1, 2n-1, n-1)$ Hadamard design and a Hadamard matrix of order $4n$.* $\square$

Chapter 3

t-designs and Steiner systems

3.1 Basic definitions and properties

In the previous chapter we considered how the 2-element subsets of the base set are contained in the blocks. More generally, we can ask the same question about t-element subsets.

Definition 3.1 *Let $S = \{1, 2, ..., v\}$. A collection $\mathcal{D}$ of distinct k-element subsets of S is called a $t-(v, k, \lambda)$ design if $0 < t \leq k < v$, $\lambda > 0$ and every t-element subset of S is contained in exactly λ of the sets in $\mathcal{D}$. A Steiner system $S(t, k, v)$ is a $t-(v, k, 1)$ design.*

In general, we call $t-(v, k, \lambda)$ designs *t-designs*.

Theorem 3.2 *If $\mathcal{D}$ is a $t-(v, k, \lambda)$ design and $0 < s < t$, then $\mathcal{D}$ is also an $s-(v, k, \lambda\binom{v-s}{t-s}/\binom{k-s}{t-s})$ design.*

Proof. Assume that $0 \leq s < t$ (zero allowed!) and denote by λ_s the number of blocks of $\mathcal{D}$ containing a given s-element subset $A \subseteq S$. We count in two ways the cardinality of the set

$$\{(C, B) \mid B \in \mathcal{D}, A \subseteq C \subseteq B, |C| = t\}.$$

There are $\binom{v-s}{t-s}$ t-element subsets C containing A, each in turn contained in exactly λ blocks $B \in \mathcal{D}$. Hence the set has cardinality $\lambda\binom{v-s}{t-s}$. On the other hand, there are exactly λ_s blocks B containing A and for each such block the intermediate set C can be chosen in $\binom{k-s}{t-s}$ ways. Hence the set has cardinality $\lambda_s\binom{k-s}{t-s}$. Consequently,

$$\lambda\binom{v-s}{t-s} = \lambda_s\binom{k-s}{t-s}$$

which shows that λ_s is independent of the choice of A and equals $\lambda\binom{v-s}{t-s}/\binom{k-s}{t-s}$. □

From the proof of Theorem 3.2 we immediately obtain the following corollary. The quantities λ_0 and λ_1 were defined in the proof of Theorem 3.2.

Corollary 3.3 *If $\mathcal{D}$ is a $t-(v,k,\lambda)$ design, there are*

$$b = \lambda_0 = \frac{\lambda\binom{v}{t}}{\binom{k}{t}} \tag{3.4}$$

blocks in $\mathcal{D}$ and every element of the base set appears in exactly

$$r = \lambda_1 = \frac{\lambda\binom{v-1}{t-1}}{\binom{k-1}{t-1}} \tag{3.5}$$

blocks. □

Notice that the right-hand sides of (3.4) and (3.5) must be integers.

Example 3.6 There exist no Steiner systems $S(4,5,9)$, because by (3.4) the number of blocks should be $\binom{9}{4}/\binom{5}{4}$, which is not an integer. □

Theorem 3.7 *If $\mathcal{D}$ is a $t-(v,k,\lambda)$ design, $t \geq 2$ and $i \in S$, then the set*

$$\mathcal{D}_i = \{B \setminus \{i\} \mid B \in \mathcal{D}, i \in B\}$$

is a $(t-1)-(v-1,k-1,\lambda)$ design.

Proof. For every $(t-1)$-element subset T of $S\setminus\{i\}$, the number of blocks of $\mathcal{D}_i$ containing T is the same as the number of blocks of $\mathcal{D}$ containing $T \cup \{i\}$.□

If there is a Steiner system $S(t,k,v)$, then by Theorem 3.7 there also exist Steiner systems $S(t-1,k-1,v-1)$, $S(t-2,k-2,v-2)$, ..., $S(1,k-t+1,v-t+1)$. By (3.4) this means all the numbers

$$\binom{v}{t}/\binom{k}{t}, \binom{v-1}{t-1}/\binom{k-1}{t-1}, \ldots, \binom{v-t+1}{1}/\binom{k-t+1}{1}$$

must be integers.

Theorem 3.8 *If a Steiner system $S(2,3,v)$ exists, then $v \equiv 1$ or $3 \pmod 6$.*

Proof. By the previous discussion $\binom{v}{2}/\binom{3}{2} = \frac{1}{6}v(v-1)$ and $\binom{v-1}{1}/\binom{2}{1} = \frac{1}{2}(v-1)$ are integers. By the former, $v(v-1) \equiv 0 \pmod 6$, i.e., $v \equiv 0,1,3$ or $4 \pmod 6$, and of these only $v \equiv 1$ or $3 \pmod 6$ are possible by the latter.□

Theorem 3.9 *If there exists a Hadamard $2-(4n-1, 2n-1, n-1)$ design, then there also exists a $3-(4n, 2n, n-1)$ design.*

Proof. Assume that $\mathcal{D}$ is a $2-(4n-1, 2n-1, n-1)$ design and $S = \{1, 2, \ldots, 4n-1\}$. We define

$$\mathcal{E} = \{S \setminus B \mid B \in \mathcal{D}\} \cup \{B \cup \{4n\} \mid B \in \mathcal{D}\}.$$

We claim that $\mathcal{E}$ is a $3-(4n, 2n, n-1)$ design. Every block of $\mathcal{E}$ contains $2n$ elements of the set $S \cup \{4n\}$. It suffices to show that every 3-element subset of $S \cup \{4n\}$ is contained in exactly $n-1$ blocks of $\mathcal{E}$.

If $a, b \in S$, $a \neq b$, then $\{a, b, 4n\} \subseteq B \cup \{4n\}$ where $B \in \mathcal{D}$ if and only if $\{a, b\} \subseteq B$, and hence $\{a, b, 4n\}$ is contained in exactly $n-1$ blocks.

Assume therefore that $a, b, c \in S$, $a \neq b \neq c \neq a$. For every $T \subseteq \{a, b, c\}$, denote the number of blocks of $\mathcal{D}$ containing T by n_T. The number of blocks $S \setminus B$ and $B \cup \{4n\}$ of $\mathcal{E}$ containing the set $\{a, b, c\}$ is $n_\emptyset + n_{\{a,b,c\}}$. By the principle of inclusion and exclusion

$$\begin{aligned}
n_\emptyset &= b - (n_{\{a\}} + n_{\{b\}} + n_{\{c\}}) + (n_{\{a,b\}} + n_{\{b,c\}} + n_{\{c,a\}}) - n_{\{a,b,c\}} \\
&= b - 3r + 3\lambda - n_{\{a,b,c\}} \\
&= 4n - 1 - 3(2n-1) + 3(n-1) - n_{\{a,b,c\}} \\
&= n - 1 - n_{\{a,b,c\}}
\end{aligned}$$

and hence $n_\emptyset + n_{\{a,b,c\}} = n - 1$. □

Together with Example 2.22 and Theorem 2.26, Theorem 3.9 shows that there exists an infinite family of 3-designs.

3.2 Steiner triple systems

Theorem 3.10 *A Steiner system $S(2,3,v)$ exists if and only if $v \equiv 1$ or $3 \pmod 6$.*

Proof. By Theorem 3.8, it is sufficient to prove the existence of $S(2,3,v)$ when $v \equiv 1,3 \pmod 6$.

Case $v = 6t+3$: Take as the base set S all the ordered pairs (i,k) where $i \in \mathbb{Z}_{2t+1}$, $k \in \mathbb{Z}_3$. Choose as the triples

$$\{(i,0),(i,1),(i,2)\} \text{ for all } i \in \mathbb{Z}_{2t+1}$$

$$\{(i,k),(j,k),((t+1)(i+j),k+1)\} \text{ for all } i,j \in \mathbb{Z}_{2t+1}, i \neq j, k \in \mathbb{Z}_3,$$

the number of which is $b = (2t+1) + 3\binom{2t+1}{2} = \frac{1}{3}\binom{v}{2}$. We show that every two different elements $(i,k),(a,b) \in S$ are contained in at least one triple. Together the triples contain $3b = \binom{v}{2}$ 2-element subsets of S, hence all of them exactly once.

If $a = i$, the two elements are contained in a triple of the first type; assume $a \neq i$. If $b = k$, the two elements are contained in a triple of the second type. If $b \neq k$, then $b = k+1$ (or $k = b+1$, a symmetric case). Now the two elements are contained in the triple of the second type with $j = 2a - i$, because $(t+1)(i+2a-i) = (2t+2)a = a$ in $\mathbb{Z}_{2t+1}$. Notice that our choice for j was legal: we did not choose $j = i$ because we got a and not i.

Case $v = 6t+1$: The Latin square $L_1(x,y) = x+y$ over $\mathbb{Z}_{2t}$ of Example 2.48 is symmetric and its main diagonal entries $L(1,1)$, $L(2,2)$, $\ldots$, $L(2t,2t)$ are $2,4,\ldots,2t,2,4,\ldots,2t$. By relabelling the entries we trivially obtain a symmetric Latin square $L(x,y)$ over $\mathbb{Z}_{2t}$ whose main diagonal entries are $1,2,\ldots,t,1,2,\ldots,t$, in that order.

Take as the elements of the base set S the symbol ∞ and all the ordered pairs (i,k), where $i \in \mathbb{Z}_{2t}$ and $k \in \mathbb{Z}_3$. Choose as the triples

$$\{(i,0),(i,1),(i,2)\}, 1 \leq i \leq t$$

$$\{(i,k),(i-t,k+1),\infty\}, t+1 \leq i \leq 2t, k \in \mathbb{Z}_3$$

$$\{(i,k),(j,k),(L(i,j),k+1)\}, i,j \in \mathbb{Z}_{2t}, i \neq j, k \in \mathbb{Z}_3,$$

the number of which is $t + 3t + 3\binom{2t}{2} = \frac{1}{3}\binom{v}{2}$. Clearly, ∞ appears with every other element of S in a unique block. It suffices to prove that any two different elements (i,k) and (a,b) of S are contained in at least one block.

If $b = k$, the two elements both appear in a block of the third type; assume $b \neq k$. By symmetry, assume $b = k+1$. Then the two elements are contained in a block of the third type if we can find $j \neq i$ such that $L(i,j) = a$. Since $L(x,y)$ is a Latin square, this is always possible, except when $a = L(i,i)$, i.e., $a = i$ for $1 \leq i \leq t$ and $a = i - t$ for $t+1 \leq i \leq 2t$; but then the two elements are contained in a block of the first or second type, respectively. $\square$

Chapter 4

Codes and designs

4.1 Basics on codes

Denote the elements of the field $\mathbb{Z}_2$ by 0 and 1. The set $\mathbb{Z}_2^n$ is called the *binary Hamming space*. Its elements are called (binary) *vectors* or (binary) *words*. If $\mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n$, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, the *Hamming distance* $d(\mathbf{x}, \mathbf{y})$ between $\mathbf{x}$ and $\mathbf{y}$ is defined to be the number of indices i such that $x_i \neq y_i$. The triangle inequality

$$d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \geq d(\mathbf{x}, \mathbf{z}) \text{ for all } \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{Z}_2^n$$

holds for the Hamming distance: we obtain $\mathbf{y}$ from $\mathbf{x}$ by changing $d(\mathbf{x}, \mathbf{y})$ coordinates in $\mathbf{x}$, and $\mathbf{z}$ from $\mathbf{y}$ by changing $d(\mathbf{y}, \mathbf{z})$ coordinates in $\mathbf{y}$; clearly $\mathbf{z}$ can therefore be obtained by changing at most $d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ coordinates in $\mathbf{x}$.

The *Hamming sphere* of radius r centred at $\mathbf{x} \in \mathbb{Z}_2^n$ is the set

$$B_r(\mathbf{x}) = \{\mathbf{y} \in \mathbb{Z}_2^n \mid d(\mathbf{y}, \mathbf{x}) \leq r\}.$$

Clearly

$$|B_r(\mathbf{x})| = \sum_{i=0}^{r} \binom{n}{i} \tag{4.1}$$

A nonempty subset of $\mathbb{Z}_2^n$ is called a *binary code C of length n*. The elements of C are called *codewords*. For convenience, we assume that codes have at least two codewords. The *minimum distance* of C is the smallest of the pairwise Hamming distances between different codewords. A code with minimum distance at least $2e + 1$ is called *e-error-correcting.*

Assume that C is an e-error-correcting code with K codewords. Then, by the triangle inequality, the Hamming spheres of radius e centred at the codewords of C are disjoint. Therefore the number cardinality K of C satisfies the *Hamming bound*

$$K \leq \frac{2^n}{\sum_{i=0}^{e} \binom{n}{i}}. \tag{4.2}$$

If equality holds, i.e., the Hamming spheres are disjoint and their union is the whole Hamming space $\mathbb{Z}_2^n$, then C is called *perfect.*

If $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}_2^n$, and $\mathbf{y} = (y_1, \ldots, y_n) \in \mathbb{Z}_2^n$, then

$$\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n),$$

and

$$\mathbf{x} * \mathbf{y} = (x_1 y_1, \ldots, x_n y_n).$$

In particular, $\mathbf{x} * \mathbf{y}$ has 1 in coordinate i if and only if both $x_i = 1$ and $y_i = 1$.

The number of 1's in $\mathbf{x} \in \mathbb{Z}_2^n$ is called the *weight* $w(\mathbf{x})$ of $\mathbf{x}$. The vector $(0, 0, \ldots, 0) \in \mathbb{Z}_2^n$ is called the *all-zero word* and is denoted by $\mathbf{0}$. Clearly,

$$d(\mathbf{x}, \mathbf{y}) = w(\mathbf{x} + \mathbf{y}) \text{ for all } \mathbf{x}, \mathbf{y} \in \mathbb{Z}_2^n. \tag{4.3}$$

We further denote

$$\langle \mathbf{x}, \mathbf{y} \rangle = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n,$$

which is always an element of $\mathbb{Z}_2$. Clearly,

$$\langle \mathbf{x} + \mathbf{y}, \mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle + \langle \mathbf{y}, \mathbf{z} \rangle, \tag{4.4}$$

where the addition on the right-hand side is performed in $\mathbb{Z}_2$.

Lemma 4.5 *$w(\mathbf{x} + \mathbf{y}) = w(\mathbf{x}) + w(\mathbf{y}) - 2w(\mathbf{x} * \mathbf{y})$.*

Proof. From the figure

$$\begin{array}{lcllll}
\mathbf{x} & = & 111\ldots1 & 000\ldots0\ 111\ldots1 & 000\ldots0 \\
\mathbf{y} & = & \underbrace{111\ldots1}_{w(\mathbf{x}*\mathbf{y})} & \underbrace{111\ldots1\ 000\ldots0}_{w(\mathbf{x}+\mathbf{y})} & 000\ldots0
\end{array}$$

we see that $w(\mathbf{x}) + w(\mathbf{y}) = w(\mathbf{x} + \mathbf{y}) + 2w(\mathbf{x} * \mathbf{y})$, as claimed. $\square$

4.2 Binary Golay code

Let $\mathbf{A}$ be the 11×11 matrix whose rows are 11011100010 and all its cyclic shifts, i.e., the words 11011100010, 01101110001, ..., 10111000101. By Example 2.56, $\mathbf{A}$ is an incidence matrix of a symmetric $2-(11,6,3)$ design, and by Theorem 2.14 any two rows have exactly three 1's in common, i.e., the Hamming distance between any two rows equals six.

Consider now the matrix

$$\mathbf{G} = \left(\begin{array}{c|c|c|c} 1 & & 0 & \\ \vdots & \mathbf{I}_{11} & \vdots & \mathbf{A} \\ 1 & & 0 & \\ \hline 0 & 00\ldots0 & 1 & 11\ldots1 \end{array} \right),$$

whose entries belong to $\mathbb{Z}_2$.

Definition 4.6 *The extended binary Golay code $\mathcal{G}_{24}$ is the set of all the 2^{12} words in $\mathbb{Z}_2^n$ that can be obtained as sums of the rows of $\mathbf{G}$.*

The number of codewords in $\mathcal{G}_{24}$ is indeed 2^{12}. Since we are dealing with $\mathbb{Z}_2^n$, the sum of a row with itself is $\mathbf{0}$, and therefore we only need to consider the sums of arbitrary subsets of rows. By convention, the sum corresponding to the empty sum is $\mathbf{0}$. Because the columns 2–13 form the identity matrix $\mathbf{I}_{12}$ it is clear that all these 2^{12} sums are different.

Lemma 4.7 *If $\mathbf{c} \in \mathcal{G}_{24}$ and $\mathbf{r}$ is a row of the matrix $\mathbf{G}$ then $w(\mathbf{c} * \mathbf{r})$ is even (or equivalently, $\langle \mathbf{c}, \mathbf{r} \rangle = 0$).*

Proof. Assume first that also $\mathbf{c}$ is a row of $\mathbf{G}$. If $\mathbf{c} = \mathbf{r}$, then $w(\mathbf{c} * \mathbf{r}) = w(\mathbf{c}) = 8$ or 12. If $\mathbf{c} \neq \mathbf{r}$ and neither of them is the last row, then $\mathbf{c}$ and $\mathbf{r}$ have 1's in common in the first column and in three columns of $\mathbf{A}$ and $w(\mathbf{c} * \mathbf{r})$ is 4. Finally, if $\mathbf{r}$ is the last row, and $\mathbf{c}$ is not, then $w(\mathbf{c} * \mathbf{r}) = 6$.

In general, $\mathbf{c} = \mathbf{r}_1 + \mathbf{r}_2 + \ldots + \mathbf{r}_k$ for some k rows $\mathbf{r}_i$ of $\mathbf{G}$. By (4.4), $\langle \mathbf{c}, \mathbf{r} \rangle = \langle \mathbf{r}_1, \mathbf{r} \rangle + \ldots \langle \mathbf{r}_k, \mathbf{r} \rangle = 0$. □

Lemma 4.8 *The weight of each codeword of $\mathcal{G}_{24}$ is divisible by 4.*

Proof. Assume that $\mathbf{c}$ is the sum of some k rows of $\mathbf{G}$. The proof is by induction on k. The claim is clear for $k = 1$. Assume that it is true whenever $\mathbf{c}$ is a sum of at most k rows of $\mathbf{G}$.

Assume that $\mathbf{c} = \mathbf{r}' + \mathbf{c}'$ where $\mathbf{r}'$ is a row of $\mathbf{G}$ and $\mathbf{c}'$ is a sum of k rows of $\mathbf{G}$. Then by Lemma 4.8

$$w(\mathbf{c}) = w(\mathbf{r}') + w(\mathbf{c}') - 2w(\mathbf{r}' * \mathbf{c}'),$$

which is divisible by four, because the first two terms are divisible by 4 by the induction hypothesis, and the last one by the previous lemma. □

If $\mathbf{c} \in \mathcal{G}_{24}$, we denote by $w_L(\mathbf{c})$ the weight of the left half, i.e., the number of 1's in the first 12 coordinates, and by $w_R(\mathbf{c})$ the weight of the right half.

Theorem 4.9 *If $\mathbf{c} \in \mathcal{G}_{24}$ and $\mathbf{c} \neq \mathbf{0}$, then $w(\mathbf{c}) \geq 8$. Consequently, if $\mathbf{x}, \mathbf{y} \in \mathcal{G}_{24}$ and $\mathbf{x} \neq \mathbf{y}$, then $d(\mathbf{x}, \mathbf{y}) \geq 8$.*

Proof. We have already shown that the weight of a codeword is divisible by four. It remains to prove that it cannot be equal to four. Clearly $w_L(\mathbf{c})$ is even for every codeword $\mathbf{c}$: if $\mathbf{c}$ is the sum of some i of the first 11 rows then $w_L(\mathbf{c})$ is i for even i and $i+1$ for odd i. Assume that $\mathbf{c} \in \mathcal{G}_{24}$ has weight four.

Case 1 $w_L(\mathbf{c}) = 0$: Then $\mathbf{c}$ is $\mathbf{0}$ or the last row of $\mathbf{G}$, neither of weight four.

Case 2 $w_L(\mathbf{c}) = 2$: Then $\mathbf{c}$ is the sum of one or two rows of $\mathbf{G}$ (possibly together with the last row), but the sum of two rows of $\mathbf{A}$ has always weight 6, and hence $w(\mathbf{c}) \geq 2 + 6 = 8$.

Case 3 $w_L(\mathbf{c}) = 4$: Then $\mathbf{c}$ is the sum of three or four of the first 11 rows; the last row cannot be involved because $w_R(\mathbf{c}) = 0$. If $\mathbf{c}$ is a sum of three rows and $\mathbf{r}$ is any other row (of the first 11 rows), then $w_R(\mathbf{c}+\mathbf{r}) = w_R(\mathbf{r}) = 6$ and $w_L(\mathbf{c}+\mathbf{r}) = 4$ because $\mathbf{c}+\mathbf{r}$ is a sum of four rows. But then $\mathbf{c}+\mathbf{r} \in \mathcal{G}_{24}$ has weight 10, contradicting Lemma 4.8. If $\mathbf{c}$ is a sum of four rows and $\mathbf{r}$ is one of them, then $\mathbf{c} = \mathbf{c}' + \mathbf{r}$ where $\mathbf{c}'$ is a sum of three rows. Then $w_L(\mathbf{c}') = 4$ and $w_R(\mathbf{c}') = w_R(\mathbf{r}) = 6$, which again gives a codeword of weight 10.

The last claim immediately follows from (4.3). □

Definition 4.10 *The binary Golay code $\mathcal{G}_{23}$ is obtained by deleting the first coordinate in each codeword of $\mathcal{G}_{24}$.*

Theorem 4.11 *The code $\mathcal{G}_{23}$ is a perfect 3-error-correcting code.*

Proof. Any two codewords in $\mathcal{G}_{23}$ still differ in at least seven coordinates, i.e., $\mathcal{G}_{23}$ is 3-error-correcting. Since $1 + 23 + \binom{23}{2} + \binom{23}{3} = 2^{11}$, the Hamming bound (4.2) holds with equality. □

4.3 Steiner systems $S(5,8,24)$ and $S(5,6,12)$

Theorem 4.12 *There are* 759 *codewords of weight* 8 *in* $\mathcal{G}_{24}$*.*

Proof. We show that there are 253 codewords of weight 7 and 506 codewords of weight 8 in $\mathcal{G}_{23}$, which clearly implies the claim.

Because $\mathcal{G}_{23}$ is perfect, every binary vector of weight 4 is contained in exactly one of the spheres $B_3(\mathbf{c})$, $\mathbf{c} \in \mathcal{G}_{23}$. The sphere $B_3(\mathbf{0})$ does not contain any of them; and the same is true for $B_3(\mathbf{c})$ whenever $w(\mathbf{c}) \geq 8$. Consequently all vectors of weight 4 are contained in the spheres $B_3(\mathbf{c})$ where $\mathbf{c} \in \mathcal{G}_{23}$ and $w(\mathbf{c}) = 7$. Moreover, the distance from of a codeword $\mathbf{c}$ of weight 7 to a vector $\mathbf{x}$ of weight 4 is three only if the 1's in $\mathbf{x}$ are in the coordinates where also $\mathbf{c}$ has 1's; hence $B_3(\mathbf{c})$ contains exactly $\binom{7}{4}$ vectors of weight 4. Therefore the number of codewords of weight 7 in $\mathcal{G}_{23}$ is $\binom{23}{4}/\binom{7}{4} = 253$.

In the same way consider vectors of weight 5. Exactly $253\binom{7}{5}$ of them are contained in the spheres $B_3(\mathbf{c})$ for the codewords $\mathbf{c}$ of weight 7, and the remaining ones must be contained in the spheres $B_3(\mathbf{c})$ for the codewords of weight 8, and each such sphere contains $\binom{8}{5}$ of them. Therefore the number of codewords of weight 8 in $\mathcal{G}_{23}$ is $(\binom{23}{5} - 253\binom{7}{5})/\binom{8}{5} = 506$. □

Theorem 4.13 *The words of weight* 8 *in the code* $\mathcal{G}_{24}$ *form a Steiner system* $S(5,8,24)$*.*

Proof. Construct a 759×24 matrix whose rows are the 759 codewords of weight 8 and interpret it as the incidence matrix of a design. The number of blocks is 759, and each block has eight elements of the set $T = \{1, 2, \ldots, 24\}$. We show that each 5-element subset of B belongs to exactly one of the blocks.

Assume that a 5-element set lies in more than one block. Then there are two rows in the matrix with at least five 1's in common. But then the Hamming distance between these two rows is at most 6, a contradiction.

On the other hand, each of these 759 blocks contains $\binom{8}{5}$ 5-element subsets, so altogether the blocks contain $759\binom{8}{5} = \binom{24}{5}$ 5-elements subsets of T, i.e., all the 5-element subsets of T. □

Theorem 4.14 *There is a Steiner system* $S(5,6,12)$*.*

Proof. Any given 5-element subset of the set $Q = \{13, 14, \ldots, 24\}$ is contained in a unique block of our $S(5,8,24)$. By Lemma 4.7 the corresponding row $\mathbf{c}$ in the incidence matrix of $S(5,8,24)$ and the last row of $\mathbf{G}$ have an even number of 1's in common. They cannot have eight 1's in common, otherwise their sum would give a codeword of weight four in $\mathcal{G}_{24}$. Hence the number of

1's in common must be six. We can therefore take as blocks of the Steiner system $S(5,6,12)$ all the sets $B \cap Q$ where $B \in S(5,8,24)$ and $|B \cap Q| = 6$. $\square$

From Theorem 3.7 we immediately obtain the following corollary.

Corollary 4.15 *There exist Steiner systems $S(4,7,23)$ and $S(4,5,11)$.* $\square$

Apart from these four the only currently known Steiner systems $S(t,k,v)$ with $k > t \geq 4$ are $S(5,6,24)$, $S(5,7,28)$, $S(5,6,48)$, $S(5,6,72)$, $S(5,6,84)$, $S(5,6,108)$, $S(5,6,132)$ and $S(5,6,168)$, and the 4-designs resulting from Theorem 3.7. It is an open problem whether or not there exists an infinite number of such Steiner systems.

Bibliography

[1] Anderson, I.: *A First Course in Combinatorial Mathematics.* Oxford: Oxford University Press, 1989 (2nd edition).

[2] Anderson, I.: *Combinatorial Designs: Construction Methods.* Chichester: Ellis Horwood, 1990. (Together with [1] the primary reference for these lecture notes.)

[3] Anderson, I.: *Combinatorial Designs and Tournaments.* Oxford: Oxford University Press, 1997.

[4] Beth, T., Jungnickel, D. and Lenz, H.: *Design Theory.* Cambridge: Cambridge University Press, 1993.

[5] Colbourn, C.J. and Dinitz, J.H.(eds.): *The CRC Handbook of Combinatorial Designs.* Boca Raton: CRC Press, 1996.

[6] Hall, M.: *Combinatorial Theory.* New York: Wiley, 1986 (2nd edition).

[7] Hardy, G.H. and Wright, E.M.: *An Introduction to the Theory of Numbers.* Oxford: Oxford University Press, 1979 (5th edition).

[8] van Lint, J.H.: *Introduction to Coding Theory.* New York: Springer, 1992 (2nd edition).

[9] van Lint, J.H. and Wilson, R.M.: *A Course in Combinatorics.* Cambridge: Cambridge University Press, 1992.

[10] MacWilliams, F.J. and Sloane, N.J.A.: *The Theory of Error-Correcting Codes.* Amsterdam: North-Holland, 1977.

[11] Ryser, H.J.: *Combinatorial Mathematics.* Carus Math. Monograph 14, Math. Assoc. of America, 1963.