Steiner systems and finite projective planes

1 Structures and Designs

We start by introducing the basics :

Definition. A *structure* $\mathcal{S}$ is two finite sets of objects called *points* and *blocks*, together with an incidence relation $\mathcal{I}$ between them.

Notation. We will often denote blocks with lowercase letters and points with uppercase letter.

Remark. For any block, there is a set of points incident to it. We will frequently identify a block with the set of points incident to it. Hence for a block y and a point P incident to y we will often write $P \in y$ instead of $P\mathcal{I}y$.

Notation. We denote the number of points of a structure $\mathcal{S}$ with v and the number of blocks with b. Furthermore for a point P and a block y of a structure $\mathcal{S}$ we denote $\langle P\rangle$, and $\langle y\rangle$, the set of blocks incidents to P and the set of points incident to y. By abuse of notation we will write $|P| := |\langle P\rangle|$ and $|y| := |\langle y\rangle|$.

We are mostly interested in finite structures. Let us give a first example :

Example. We define the Structure $\mathcal{S}_1$, with points $\{1,2,3,4,5\}$ and blocks :

$$a_1 = \{1,2,3\},\ a_2 = \{1,2,3\},\ a_3 = \{2,3,4\},\ a_4 = \{3,4,5\}$$

Here 1 would be incident to a_1, a_2, and 5 would be incident to a_4

Notice that in the example above, we have a block that is repeated, this leads to the following construction :

Definition. We define the equivalence relation $\mathcal{R}$ on the set of blocks of a structure $\mathcal{S}$, through $x\mathcal{R}y$ if $\langle x\rangle = \langle y\rangle$ (*i.e.* both of the blocks are incident to the same points). A block whose equivalence class has at least 2 elements is called a *repeated block*. We now define the *reduction* of $\mathcal{S}$, the structure $\mathcal{S}/\mathcal{R}$, with the same points as $\mathcal{S}$, and whose blocks are the equivalence classes defined above. A structure without any repeated blocks is called *reduced*.

Definition. For a structure $\mathcal{S}$ with blocks $y_1, \dots, y_b$, and points $P_1, \dots, P_v$ we define the vxb *incidence matrix* A of $\mathcal{S}$ through :

$$a_{i,j} = \begin{cases} 1 \text{ if } P_i \in x_j \\ 0 \text{ else} \end{cases}$$

Remark. Notice that a structure is reduced, if and only if no two columns of the incidence matrix are identical.

Example. In the example above the incidence matrix of $\mathcal{S}_1$ is :

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

Again we notice that the first and second columns are identical which corresponds to our repeated block.

Remark. Notice that the incidence matrix of a structure, is not unique it depends on the indexing of the points and blocks.

We will also give a more 'daily life' example of a structure, which will come in handy latter :

Example. Take $\mathbb{R}^2$ and consider a set of points and of lines. The lines can be seen as blocks, as they are incident to certain points.

Definition. A structure $\mathcal{S}$ is called *uniform* if it has at least one block and each block contains exactly $k > 0$ points.

Definition. A *design* is a uniform, reduced structure. This means that it has no repeated blocks, and each block has the same number of points.

Definition. Let $\mathcal{S}$ be a structure, with $v > 0$ points. Assume that there exist integers λ, t with $\lambda > 0$ and $0 \leq t \leq v$, and such that every subset of t points of $\mathcal{S}$ is incident with exactly λ blocks. We call such a structure a *t-structure for λ*, or simply a *t-structure*.

Definition. A uniform t-structure with block size k is called is called a $t-(v,k,\lambda)$ *structure*. Here v is as always the number of points, and λ is the number of common blocks on t points.

Example. Consider the structure $\mathcal{S}_2$ with points $\{1,2,3,4,5,6\}$ and with blocks :

$$d_1 = \{1,2,3\}, d_2 = \{4,5,6\}, d_3 = \{1,4\}, d_4 = \{1,5\}, d_5 = \{1,6\}, d_6 = \{2,4\},$$

$$d_7 = \{2,5\}, d_8 = \{2,6\}, d_9 = \{3,4\}, d_{10} = \{3,5\}, d_{11} = \{3,6\}$$

We can check that S_2 is a 2-structure with $\lambda = 1$, and a 1-structure with $r = 4$.

Now it seems the next logical step would be to combine designs and t-structures, which we will do next.

Definition. $\mathcal{S}$ is a *t-design* if it is a t-structure which is also a design. Instead of saying that $\mathcal{S}$ is a t-design for (v,k,λ), we simply say $\mathcal{S}$ is a t-(v,k,λ).

2 Steiner Systems, definition, existence

Definition. A *Steiner System* is a $t-(v,k,\lambda)$ design with $\lambda = 1$, that is every subset of t points is contained in exactly one block. We write $S(t,k,v)$, for this Steiner system.

Remark. For a set of points $X = \{1,\dots,v\}$, we can represent each block b, through its indicator (characteristic vector), $(c_1,\dots,c_v)$, where $c_i = 1$ if $i \in b$ and $c_i = 0$, else. Notice that these are simply the columns of the incidence matrix.
In this way the a t-design become a binary code of length v, in which every codeword has weight k.

Example. Consider the structure with points :

$$\{001, 010, 011, 100, 101, 110, 111\}$$

and with blocks :

$$\begin{aligned} b_1 &= \{001, 011, 010\}, b_5 = \{010, 101, 111\} \\ b_2 &= \{001, 100, 101\}, b_6 = \{100, 011, 111\} \\ b_3 &= \{010, 100, 110\}, b_7 = \{011, 101, 110\} \\ b_4 &= \{001, 110, 111\} \end{aligned} \tag{1}$$

Notice that then we have a Steiner system $S(2,3,7)$. That is we have blocks of size 3, and every subset of 2 points is contained in exactly one block ($\lambda = 1$).

Through combinatorial arguments we have that the parameters of a t-design must satisfy the following :

Theorem. ***(i)*** *For distinct points $P_1,\dots,P_t$ of a t-design, and for $1 \le i \le t$, we have that the number of blocks containing $P_1,\dots,P_i$ is :*

$$\lambda_i = \lambda \frac{\binom{v-i}{t-i}}{\binom{k-i}{t-i}}$$

(ii) *Furthermore the total number of blocks must be :*

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

(iii) *Finally each point must belong to r blocks (by definition of t-design). And we have :*

$$bk = vr \text{ and } \lambda_2(v-1) = r(k-1)$$

Proof. We will only prove the first statement, as the others follow from this one. Let $\mathcal{B}$ be a subset of i points of $\mathcal{S}$. Let λ_i be the number of blocks which contain $\mathcal{B}$. We will say that $(\mathcal{T}, y)$ is an admissible pair, for a set $\mathcal{T}$ of t points which contain B and a block y which contains $\mathcal{T}$.
Each of the λ_i blocks which contain $\mathcal{B}$ contains $\binom{k-s}{t-s}$ t-sets which contain $\mathcal{B}$. So there are $\lambda_i\binom{k-s}{t-s}$ admissible pairs. There are $\binom{v-s}{t-s}$ ways to choose a set $\mathcal{T}$ of t points which contains $\mathcal{B}$, and because

we have a t-design each of these sets is on exactly λ blocks. The number of admissible pairs is thus $\lambda\binom{v-s}{t-s}$. Now we have :

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

The result then follows immediately. ■

These conditions are necessary for the existence of a $t-(v,k,\lambda)$ design, are they sufficient ?

Definition. The *divisibility condition* is that for every i such that $0 \leq i \leq t-1$ we have that $\binom{k-i}{t-i}$ divides $\lambda\binom{v-i}{t-i}$.

Remark. The divisibility condition is simply a rephrasing of condition **(i)** in the theorem above.

Remark. For Steiner Systems, the divisibility condition reduces to: for every i such that $0 \leq i \leq t-1$ we have that $\binom{k-i}{t-i}$ divides $\binom{v-i}{t-i}$.

Theorem (Keevash, 2019). *The divisibility conditions are a sufficient condition for the existence of a $t-(v,k,\lambda)$ design for v large enough. In particular for v large enough, there will always exist a Steiner System $S(t,k,v)$.*

Remark. In particular for fixed λ, t, k, there are only finitely many v such that there is no $t-(v,k,\lambda)$.

This powerful theorem can be used to prove statements such as the following :

Example. Steiner-system $S(2,3,v)$ exists if and only if $v \equiv 1,3 \mod 6$.
We must check that the divisibility condition holds : $\binom{3-i}{2-i}$ must divide $\binom{v-i}{2-i}$, for all $i \leq t-1 = 1$, hence for $i = 0, 1$, that is : $3-i$ must divide $\frac{(v-i)!}{(v-2)!(2-i)!}$. For $i = 0$, this equates to 3 must divide $\binom{v}{2}$, hence v must be a multiple of 3 or $v-1$ must be a multiple of 3, *i.e.* $v \equiv 0,1 \mod 3$.
A similar statement holds for $i = 1$, that is 2 must divide $\binom{v-1}{1}$, hence 2 must divide $v-1$, that is $v \equiv 1 \mod 2$.
Combining these two results we get that : $v \equiv 1,3 \mod 6$, by Chinese remainder theorem. By Keevash this also means that it is a sufficient condition for the existence of $S(2,3,v)$ systems.

3 Finite Projective Planes

Definition. A *projective geometry* $\mathcal{P}$ is a structure of *points* and *lines* (what we called blocks until now), such that :

- **(P1)** every pair of points are on a unique common line
- **(P2)** every line contains at least 3 points
- **(P3)** $\mathcal{P}$ contains a set of 3 points which are not on a common line.
- **(P4)** If P_1, P_2, P_3 are distinct points and l_1, l_2, l_3 are distinct lines with P_i on l_j for all $i \neq j$, and if l is a line intersecting l_2 and l_3 but not containing P_1, then l intersects l_1

Furthermore we say that a *projective plane* is a projective geometry satifying :

(P5) every pair of distinct lines contain a common point.

Remark. Intuitively the fourth condition means : if a line intersects two sides of a triangle but does not contain their common vertex, then it intersects the third side. Notice that in the case of projective planes, the fifth condition together with the first one implies the fourth one.

Definition. A *finite projective plane* is simply a projective plane with a finite number of points.

Proposition. *A finite projective plane of order n is a $2-(n^2+n+1, n+1, 1)$ design, which means it is a Steiner system $S(2, n+1, n^2+n+1)$. Intuitively this means that we have n lines (blocks), n^2+n+1 points, and that every pairs of two points is contained in exactly one ($\lambda = 1$) line.*

Remark. We know of finite projective planes for all prime power orders, we will give an explicit construction below. For non-prime power order, we do not know of any finite projective plane.

Construction (Desarguesian planes). *Let $\mathbb{F} = \mathbb{F}_q$ be a finite field, for q a prime power. Now let V be an $(n+1)$-dimensional vector space over $\mathbb{F}$. The projective space, or projective geometry $PG(n, q)$ is the set of all vector subspace of V. Notice that it is partially ordered with respect to set-theoretic inclusion (i.e. a line will be contained in a plane, a plane in a hyper-plane, and so on).*
An i-flat is a subspace of the vector-space of dimension $i+1$. We will call 0-flats points, 1-flats lines, 2-flats planes and $(n-1)$-flats will be called hyper-planes.
We can identify any flat with the points (0-flats) that it contains. Now notice the following : for i such that $1 \leq i \leq n-1$, we have that the points and i-flats form a 2-design. In particular the points (0-flats) and lines (1-flats) form a $2-(\frac{q^{n+1}-1}{q-1}, q+1, 1)$; notice that this is a Steiner system $S(2, q+1, \frac{q^{n+1}-1}{q-1})$. This is because every set of 2 points, lies on exactly one line.
We can do something similar for the set of points and of hyperplanes ($(n-1)$-flats). We then get a $2-(\frac{q^{n+1}-1}{q-1}, \frac{q^n-1}{q-1}, \frac{q^{n-1}-1}{q-1})$ design.
Now for $n = 2$, these two cases collapse and we get $PG(2, q)$ which by the above construction is a $2-(q^2+q+1, q+1, 1)$ design, i.e. a Steiner system $S(2, q+1, q^2+q+1)$.

Remark. The construction above relies on the existence of a finite field $\mathbb{F}$, which only exists in cases where the order q is a prime power. This is why this construction only holds for such q.

4 The Bruck-Ryser-Chowla Theorem

Definition. If a structure has as many points as blocks, we call it a *square structure*.

Remark. The definition above is equivalent to $v = b$ and to the fact that the incidence matrix of the structure is a square matrix.

Definition. A *symmetric design* is a square $2-(v, k, \lambda)$ design, with $v > k$.

Remark. In a symetric (v, k, λ) design, $b = v$ and $r = k$. From the theorem in section 2, it then follows that :

$$(v-1)\lambda = k(k-1) \text{ hence } v = \frac{k(k-1)+\lambda}{\lambda}$$

Integers v, k, λ satisfying this relation, do not guarantee the existence of a symmetric $2-(v, k, \lambda)$. We must further restrict our assumptions which leads to the following theorem :

Theorem (Bruck-Ryser-Chowla). *If v, k, λ are integers satisfying $(v-1)\lambda = k(k-1)$, then for the existence of a symmetric $2-(v,k,\lambda)$ it is necessary that :*

(i) if v is even then $k-\lambda$ is a square

(ii) if v is odd then $z^2 = (k-\lambda)x^2 + (-1)^{\frac{v-1}{2}}\lambda y^2$ has non-trivial solution in integer x, y, z.

We will only prove the case for v odd. The proof reacquires some results from number theory, summarized here :

Lemma (1). *Every positive integer is the sum of four integral squares.*

Lemma (2). *The following identity holds :*

$$(b_1^2+b_2^2+b_3^2+b_4^2)(x_1^2+x_2^2+x_3^2+x_4^2) = y_1^2+y_2^2+y_3^2+y_4^2$$

where y_1, y_2, y_3, y_4 are defined as follows :

$$\begin{aligned} y_1 &= b_1x_1 - b_2x_2 - b_3x_3 - b_4x_4 \\ y_2 &= b_2x_1 + b_1x_2 - b_4x_3 + b_3x_4 \\ y_3 &= b_3x_1 + b_4x_2 + b_1x_3 - b_2x_4 \\ y_4 &= b_4x_1 - b_3x_2 + b_2x_3 + b_1x_4 \end{aligned} \tag{2}$$

Proof. Let $\mathcal{D}$ be a symmetric $2-(v,k,\lambda)$ design and write $n = k-\lambda$. Label the points and the blocks of $\mathcal{D}$, through $P_1, \dots, P_v$ and $c_1, \dots, c_v$, by symmetry there is the same number of both. As defined previously let $A = (a_{ij})_{i,j=1}^v$ be the incidence matrix.
Now define independent variables $x_1, \dots, x_v$, and define the v linear forms :

$$L_i := \sum_{j=1}^{v} a_{ij}x_j$$

Now consider $Q := L_1^2 + \cdots + L_v^2$. Notice that each x_i occurs in exactly k distinct L_j. Hence each x_j^2 will occur exactly k times in Q.Now how many time will x_ix_j occur for $i \neq j$? This will happen 2λ times. Both of these results together yield :

$$Q = k(x_1^2 + \cdots + x_v^2) + 2\lambda \sum_{1 \leq i < j \leq v} x_ix_j$$

which through some rearranging yields :

$$Q = L_1^2 + \cdots + L_v^2 = n(x_1^2 + \cdots + x_v^2) + \lambda(x_1 + \cdots + x_v)^2$$

Now notice that because of Lemma 1, we can rewrite n as $n = b_1^2 + b_2^2 + b_3^2 + b_4^2$. Then we get, by Lemma 2 :

$$n(x_i^2 + x_{i+1}^2 + x_{i+2}^2 + x_{i+3}^2) = y_i^2 + y_{i+1}^2 + y_{i+2}^2 + y_{i+3}^2$$

Now we assumed v to be odd. We will first assume that $v \equiv 1 \mod 4$, and then we will take care of the case $v \equiv 3 \mod 4$.

For $v \equiv 1 \mod 4$, we can apply our previously proven identity to groups of 4 at the time and we get by defining $w = x_1 + \cdots + x_v$:

$$Q = y_1^2 + \cdots + y_{v-1}^2 + nx_v^2 + \lambda w^2$$

Notice that this identity holds, for any rational forms. Now by using Lemma 2, and because the transformation from the x_i to the y_i is non-singular, we can solve the x_i in term of the ys and hence we can express w in terms of $y_1, \ldots, y_{v-1}, x_v$, and we get an identity in the independent variables $y_1, \ldots, y_{v-1}, x_v$, where the Ls and w are rational linear forms in these variables.
Because our identity holds for any linear form, it holds for any choice of y_i, hence by setting $L_1 = y_1$, if L_1 does not have the coefficient $+1$ for y_1, and setting $L_1 = -y_1$ in all other cases, we get, with help of the identity $L_1^1 = y_1^2$, and by solving y_1 as a rational linear combination of $y_2, \ldots, y_{v-1}, x_v$, the following identity in $y_2 + \cdots + y_{v-1} + x_v$:

$$L_2^2 + \cdots + L_v^2 = y_2^2 + \cdots + y_{v-1}^2 + nx_v^2 + \lambda w^2$$

We repeat this process for $L_2, \ldots, L_{v-1}$ and finally get :

$$L_v^2 = nx_v^2 + \lambda w^2$$

with x_v an independant varible, and L_v, and w rational multiples of x_v. Take x_v to be a commun multiple of all denominators appearing in L_v and w, and we get the equation in integers :

$$z^2 = (k - \lambda)x^2 + \lambda y^2$$

The case $v \equiv 3 \mod 4$, can be handled in the following way :
We once again start with :

$$L_1^2 + \cdots + L_v^2 = n(x_1^2 + \cdots + x_v^2) + \lambda(x_1 + \cdots + x_v)^2$$

By adding nx_{v+1}^2, for a new independent variable x_{v+1}, to both sides we get :

$$L_1^2 + \cdots + L_v^2 + nx_{v+1}^2 = n(x_1^2 + \cdots + x_v^2 + x_{v+1}^2) + \lambda(x_1 + \cdots + x_v)^2$$

We can now apply the same method as before to get :

$$L_1^2 + \cdots + L_v^2 + nx_{v+1}^2 = y_1^2 + \cdots + y_{v+1}^2 + \lambda w^2$$

Now reduce this as before to get :

$$nx_{v+1}^2 = y_{v+1}^2 + \lambda w^2$$

With as before y_{v+1} and w rational multiples of the variable x_{v+1}, in this case this yields integer solutions to $(k - \lambda)x^2 = z^2 + \lambda y^2$. Now simply observe that $(-1)^{\frac{v-1}{2}} = -1$ to conclude. ■

Corollary. *If a symmetric design with $\lambda = 1$ and order n exists and if $n \equiv 1$ or $2 \mod 4$ then n can be expressed as the sum of two integral squares.*

Proof. Let $\mathcal{D}$ be a symmetric $2 - (v, k, 1)$ design and write $n = k - 1$. Then $v = n^2 + n + 1 = n(n + 1) + 1$, which is always odd. If $v \equiv 1$ or $2 \mod 4$ then $n \equiv 0$ or $3 \mod 4$.
If $v \equiv 1 \mod 4$ then $(-1)^{\frac{v-1}{2}} = 1$, so by Bruck-Ryser-Chowla we are looking for a solution of $nx^2 = y^2 - z^2$. Notice that $x = 0, y = z = 1$ is a non-trivial solution hence all values of k are possible. If $v \equiv 3 \mod 4$ we are looking for solutions to : $nx^2 = y^2 + z^2$. This is equivalent to $n = a^2 + b^2$, for a and b, rational numbers. An integer is the sum of two rationals squares if and only if it is the sum of two integral squares, which concludes the proof. ■

Example. There is no symmetric design with $\lambda = 1$ and $n = 6$. For instance there is no symmetric $2 - (43, 7, 1)$ design. This is because $n = 6$ cannot be written as the sum of 2 squares.