CHAPTER 9

Block Designs

9.1. Gaussian Binomial Coefficients

Let V be an n-dimensional vector space over the finite field $\mathbf{F}_q$ of q elements. We would like to determine the number of subspaces of dimension k. For example, the number of 1-dimensional subspaces is easily found as these are subspaces spanned by one element. Such an element must be non-zero and there are $q^n - 1$ ways of choosing such an element. But for each choice, any non-zero scalar multiple of it will generate the same subspace as there are $q - 1$ such multiples for any fixed vector, we get a final tally of

$$\frac{q^n - 1}{q - 1}$$

for the number of 1-dimensional subspaces of V. This gives us a clue of how to determine the general formula.

Each subspace of dimension k has a basis of k elements. Let us first count in how many ways we can write down a basis for a k-dimensional subspace of V. For the first vector, we have $q^n - 1$ choices. For the second, we have $q^n - q$ choices since we must not pick any scalar multiple of the first vector chosen. For the third vector, we have $q^n - q^2$ such vectors since we should not pick any linear combination of the first two chosen. In this way, we see that the number of ways of writing down a basis for a k-dimensional subspace is

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

On the other hand, any k-dimensional subspace is isomorphic to $\mathbf{F}_q^k$ and the number of bases it has correspond to the number of $k \times k$ non-singular matrices over $\mathbf{F}_q$. This number is easily seen to be

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

Therefore, we obtain:

S. M. Cioabă and M. R. Murty, *A First Course in Graph Theory and Combinatorics*

THEOREM 9.1.1. *Let V be a vector space of dimension n over $\mathbf{F}_q$. The number of k-dimensional subspaces in V is*

$$\binom{n}{k}_q := \frac{(q^n-1)(q^n-q)\cdots(q^n-q^{k-1})}{(q^k-1)(q^k-q)\cdots(q^k-q^{k-1})}.$$

REMARK 9.1.2. We refer to the numbers enumerated in the theorem as the q-binomial coefficients or sometimes as the Gaussian binomial coefficients. The reason for this will become apparent as we proceed. But for now, let us observe that if we think of q as a real number and take limits as $q \to 1^+$, we obtain by l'Hôspital's rule that

$$\lim_{q\to 1^+} \binom{n}{k}_q = \binom{n}{k},$$

and for this reason (and others), these numbers have properties similar to the binomial coefficients. This perspective has proved useful in trying to obtain q-analogs of classical binomial identities and to understand their meaning from the standpoint of these subspaces.

Let us observe that we could have done this count in another way. Indeed, to any ordered basis, we can associate a $k \times n$ matrix with the basis vectors being the rows. We can view our subspace of dimension k as the row span of this matrix. The row span is unchanged if we perform "row operations" on it as follows. We can multiply any row by a non-zero scalar. We can add one row to another. We can interchange rows. This allows us to speak about the reduced row echelon form of a matrix. This form is characterized by the fact that the first non-zero entry of each row is a 1. For any row, all the entries preceding the leading 1 are zero. If a column contains a leading 1, then all its other entries are zero. For example, if $n = 4$ and $k = 2$, the possible echelon forms are given by

$$\begin{pmatrix} 1 & 0 & * & * \\ 0 & 1 & * & * \end{pmatrix}, \quad \begin{pmatrix} 1 & * & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix}, \quad \begin{pmatrix} 1 & * & * & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

$$\begin{pmatrix} 0 & 1 & 0 & * \\ 0 & 0 & 1 & * \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & * & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

where $*$ denotes any element of $\mathbf{F}_q$. It is clear that every subspace of dimension k has a unique echelon form. Thus, the number of subspaces of dimension k is equal to the number of echelon forms for a $k \times n$ matrix over $\mathbf{F}_q$. In the above example, this number is easily seen to be

$$q^4 + q^3 + 2q^2 + q + 1 = \frac{(q^4-1)(q^4-q)}{(q^2-1)(q^2-q)}.$$

We now establish the q-analog of Pascal's triangle.

THEOREM 9.1.3.

$$\binom{n+1}{k}_q = \binom{n}{k-1}_q + q^k \binom{n}{k}_q.$$

PROOF. We prove this by counting the number of reduced row echelon forms. The left hand side is the number of reduced row echelon forms of a $k \times (n+1)$ matrix over $\mathbf{F}_q$. Such an echelon form either has a leading 1 in the $(k, n+1)$-entry or it does not. For those that do, we see that the $(k-1) \times n$ matrix formed by the first $k-1$ rows and first n columns is in echelon form and their number is

$$\binom{n}{k-1}_q.$$

If the $(k, n+1)$ entry is not a leading 1, then the last column of such a reduced row echelon form has arbitrary entries. The $k \times n$ submatrix obtained by taking the first n columns is in reduced row echelon form and thus counts the number of subspaces of dimension k in an n-dimensional vector space. This number is

$$\binom{n}{k}_q.$$

As we have q^k choices for the last column, we obtain

$$\binom{n+1}{k}_q = \binom{n}{k-1}_q + q^k \binom{n}{k}_q.$$

This completes the proof. ∎

Note that this reduces to the usual recurrence relation for binomial coefficients when $q = 1$.

THEOREM 9.1.4.

$$\binom{n}{k}_q = \binom{n}{n-k}_q.$$

PROOF. This follows by observing that there is a bijection between the k-dimensional subspaces and the $n-k$-dimensional subspaces of the dual space. This can also be verified directly as follows. Note that

$$\binom{n}{k}_q = \frac{(q^n-1)(q^{n-1}-1)\cdots(q-1)}{(q^k-1)(q^{k-1}-1)\cdots(q-1)(q^{n-k}-1)(q^{n-k-1}-1)\cdots(q-1)}$$

which is clearly symmetric under the map $k \mapsto n-k$. ∎

By applying Theorem 9.1.3, we deduce another recurrence:

$$\binom{n+1}{k}_q = \binom{n}{k}_q + q^{n+1-k} \binom{n}{k-1}_q.$$

We will use this to prove:

THEOREM 9.1.5 (The q-binomial theorem). *For $n \geq 1$,*

$$\prod_{i=0}^{n-1}(1+q^i t) = \sum_{k=0}^{n} \binom{n}{k}_q q^{\binom{k}{2}} t^k.$$

PROOF. We use induction on n. For $n = 1$, both sides of the equation at $1 + t$. Suppose that the result is true for n. Then,

$$\prod_{i=0}^{n}(1+q^i t) = (1+q^n t)\left(\sum_{k=0}^{n} \binom{n}{k}_q q^{\binom{k}{2}} t^k\right).$$

The coefficient of t^k on the right is

$$q^{\binom{k}{2}} \binom{n}{k}_q + q^{\binom{k-1}{2}} \binom{n}{k-1}_q q^n$$

which is equal to

$$q^{\binom{k}{2}} \left(\binom{n}{k}_q + q^{n-k+1} \binom{n}{k-1}_q\right) = q^{\binom{k}{2}} \binom{n+1}{k}_q,$$

as desired. ∎

9.2. Introduction to Designs

Design theory has its origin in statistics where one must set up experiments or "clinical trials" to test the reliability of a product. Consider the following problem. Suppose that we have 7 volunteers to test 7 products. Each person is willing to test 3 products and each product should be tested by 3 people to ensure objectivity. Can we arrange the experiment so that any two people would have tested precisely one product in common?

Surprisingly, a solution is provided to this problem by the *Fano plane* (see Figure 9.1). This name honors Gino Fano (1871-1952) who was one of the pioneers of projective geometry. Consider the triangle of three points; we join each vertex to midpoint of the opposite side. The three midpoints are then joined by a circle. In this way, we have 7 points and 7 "lines". Each line would represent a product and the three vertices on a line would mark out three volunteers to test that particular product. Since any two points determine a unique line, we deduce that any two people test precisely one product in common. Observe that in this situation, we have by "duality" that any two products are simultaneously tested by precisely one person.

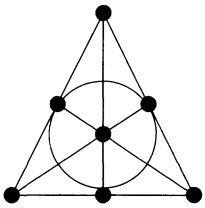

FIGURE 9.1. The Fano plane

Here is another famous problem, called Kirkman's schoolgirls problem. Thomas Kirkman (1806-1895) published this problem in *Lady's and Gentleman's Diary* in 1850. Fifteen schoolgirls walk home each day in five groups of three. Is it possible to arrange the walks over a one week period so that any two girls walk precisely once together in a group. Here is a solution. Consider the vector space $\mathbf{F}_2^4$ and remove the zero vector. We then have 15 vectors. Consider triples of vectors $\{x, y, z\}$ such that $x + y + z = 0$. The number of such vectors is 35 since we have 15 choices for x, 14 choices for y and then z is uniquely determined. Note that necessarily, these are distinct triples since if two of them were equal, we get the other vector must be zero, which we have removed. The number of ordered triples is 15×14 and we must divide this number by $3! = 6$ to get 35. Each triple corresponds to a 2-dimensional vector space of $\mathbf{F}_2^4$. It is now possible to arrange the solution vectors in 7 groups so that in each group, we have 5 triples and the union of the triples is the set of fifteen vectors. Thus, if we think each schoolgirl corresponding to a vector, this configuration gives us the solution.

To understand precisely what is behind this solution, we must understand the theory of combinatorial designs. It might be more illuminating to consider the following set up. Let X be a set of v volunteers, B a set of b products or "blocks" as they are called in the theory. We require that each volunteer test r products and each product should be tested by k people. In addition, we require that any pair of people together test precisely λ products. Can such an experiment be arranged?

We can represent this situation by a bipartite graph (X, B), where X consists of the set of v volunteers, B the set of b blocks. We join a vertex of X to a vertex of B if the corresponding person is to test that particular product. The conditions tell us that the degree of every vertex in X is r, and the degree of every vertex in B is k. The final condition tells us that any pair of vertices of X have precisely λ common neighbors. We can get immediately some necessary conditions for such

a configuration to exist. Indeed, we can count the number of edges by going through the vertices of X or by going through the vertices of B. We deduce that

$$vr = bk.$$

Now let us construct another bipartite graph in which the vertices are pairs of vertices. We join a pair to a block if they occur in that block. This gives $v(v-1)\lambda/2$ edges. Since each block has k elements in it, there are $k(k-1)/2$ pairs that each block will be joined to and so we get

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

Since $vr = bk$, we obtain

$$(v-1)\lambda = (k-1)r.$$

These are obviously necessary conditions, but they are not sufficient, as we shall see. If there is a bipartite graph satisfying these properties, we call it a $2-(v,k,\lambda)$ design. Sometimes, the more cumbersome notation of a (b,v,r,k,λ) design is used, but since v,λ and k give us r and then b by the above relations, it is prudent to drop the extra parameters. Thus, we have proved:

THEOREM 9.2.1. *In any $2-(v,k,\lambda)$ design, with b blocks and each object appearing in r blocks, we must have*

$$vr = bk, \qquad \text{and} \qquad (v-1)\lambda = (k-1)r.$$

These conditions are necessary, but as we shall see below, they are not sufficient. For instance, it will be seen that there is no way to arrange 22 objects into 22 blocks with each object occurring in precisely 7 blocks and each block containing 7 objects so that any two distinct objects occur in precisely 2 blocks. This corresponds to $(v,b,r,k,\lambda) = (22,22,7,7,2)$ or a $2-(22,7,2)$ design.

More generally, one speaks of a $t-(v,k,\lambda_t)$ design if we insist that any t points are contained in precisely λ_t blocks. For example, in the design of statistical experiments, we may want any collection t people to simultaneously test precisely λ_t products. A $2-(v,3,1)$ design is often called a *Steiner triple system*. We present examples of designs in the following sections.

9.3. Incidence Matrices

A convenient way of encoding the information in a block design (X,B) is by the use of the **incidence matrix**. This is a $v \times b$ matrix A whose rows index the objects and the columns index the blocks. The (i,j)-th entry of A is 1 if the i-th object occurs in block j. Otherwise,

it is zero. We immediately see that every row adds up to r and every column adds up to k. Also note that if we look at the $v \times v$ matrix AA^t, the (i,j)-th entry is precisely the number of common neighbors of objects i and j. By the conditions for the block design, this number is λ if $i \neq j$ and r if $i = j$. This we record as:

THEOREM 9.3.1. *Let A be the incidence matrix of the $2-(v,k,\lambda)$ block design (X,B). Let J be the $v \times v$ matrix all of whose entries are 1. Then,*

$$AA^t = \lambda J + (r-\lambda)I.$$

This relation allows us to obtain further necessary conditions for the existence of block designs. Indeed, we can compute the determinant of AA^t as

$$\begin{vmatrix} r & \lambda & \cdots & \lambda \\ \lambda & r & \cdots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \cdots & r \end{vmatrix} = \begin{vmatrix} r+(v-1)\lambda & r+(v-1)\lambda & \cdots & r+(v-1)\lambda \\ \lambda & r & \cdots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \cdots & r \end{vmatrix}$$

where we have simply added to the first row the sum of all the other rows. We can now factor $(r+(v-1)\lambda)$ from the determinant. Thus, the determinant is

$$(r+(v-1)\lambda)\begin{vmatrix} 1 & 1 & \cdots & 1 \\ \lambda & r & \cdots & \lambda \\ \vdots & \vdots & \ddots & \vdots \\ \lambda & \lambda & \cdots & r \end{vmatrix} = rk \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 0 & r-\lambda & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r-\lambda \end{vmatrix}$$
$$= rk(r-\lambda)^{v-1},$$

where we have used Theorem 9.2.1 to replace $r+(v-1)\lambda$ with rk and in the determinant we have multiplied the first row by $-\lambda$ and added it to each of the other rows. This gives $rk(r-\lambda)^{v-1}$ as the value of the determinant.

COROLLARY 9.3.2 (Fisher's inequality). *In any $2-(v,k,\lambda)$ design, we must have $b \geq v$. That is, there must be at least as many blocks as points.*

PROOF. By the theorem, we see that the matrix AA^t is non-singular and thus has rank v. If $b < v$, then as the row rank of A is equal to the column rank of A, we see that A has rank at most b. Recall that for any two matrices A and B for which AB is defined, the row space of AB is contained in the row space of A. Thus, rank of AB is less than or equal to the rank of A. In our situation, we deduce that rank of AA^t is less than or equal to b which is strictly less than v, contradiction. ∎

Designs in which $b = v$ are called *symmetric designs.* In that case, we immediately deduce:

COROLLARY 9.3.3. *If in a symmetric* $2-(v,k,\lambda)$ *design,* v *is even, then* $k-\lambda$ *is a perfect square.*

PROOF. If $b = v$, the incidence matrix is a square matrix and from the theorem, we deduce that

$$(\det A)^2 = r^2(r-\lambda)^{v-1}.$$

The left hand side is a perfect square and so $(r-\lambda)^{v-1} = (k-\lambda)^{v-1}$ must also be a perfect square. As $v-1$ is odd, this forces $k-\lambda$ to be a perfect square. ∎

Thus, in the example above, we see that there is no $2-(22,7,2)$ design because $7-2$ is not a perfect square. We will prove later the following important theorem in the theory of designs. This was proved in 1951 by Richard Hubert Bruck (1914-1991), Sarvadaman Chowla (1907-1995) and Herbert John Ryser (1923-1985).

THEOREM 9.3.4 (Bruck-Ryser-Chowla). *If* (X,B) *is a symmetric* $2-(v,k,\lambda)$ *design, and* v *is odd, then the equation*

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

has a non-zero solution in integers.

As an application of this theorem, consider the existence of a $2-(29,8,2)$ design. That is, can we arrange 29 objects into 29 blocks with each object occurring in 8 blocks and any two objects occur in precisely 2 blocks. The theorem implies that if such a design exists then we can solve the diophantine equation

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

with $(x,y,z) \neq (0,0,0)$. We may assume that $\gcd(x,y,z)=1$, for otherwise, we can cancel the common factor. From the equation, we see that 2 divides the left hand side and hence must divide the right hand side. So write $z = 2z_1$. We get

$$3x^2 + y^2 = 2z_1^2$$

has a non-trivial solution. If we reduce this mod 3, we get

$$2z_1^2 \equiv y^2 \pmod 3.$$

If z_1 is coprime to 3, we deduce that 2 is a square mod 3, which is not the case. Thus, 3 divides z_1, so write $z_1 = 3z_2$ to deduce that

$$3x^2 + y^2 = 18z_2^2$$

has a non-trivial solution. But now, 3 divides y and $9|3x^2$ implies $3|x$, contrary to the coprimality assumption at the outset. Hence, there is no such design.

9.4. Examples of Designs

If we consider a v element set X and consider the collection B of all k-element subsets of X, we see that any t-element set with $0 \leq t \leq k$, is contained in precisely

$$\binom{v-t}{k-t}$$

elements of B. This is an example of a

$$t - \left(v, k, \binom{v-t}{k-t}\right)$$

design.

We will now consider q-analogs of this construction. We begin with an important class of examples known as *projective planes.* For the elements of X we take all the 1-dimensional subspaces of $V = \mathbf{F}_q^3$. There are

$$\binom{3}{1}_q = \frac{q^3-1}{q-1} = q^2+q+1$$

such subspaces. For B we take the 2-dimensional subspaces and we will say a 1-dimensional subspace U is incident with a two dimensional subspace W if $U \subseteq W$. By the correspondence theorem, the number of such subspaces is the same as the number of 1-dimensional subspaces of the quotient V/U. As this quotient is isomorphic to $\mathbf{F}_q^2$, the number of times a subspace is replicated in the blocks is $(q^2-1)/(q-1) = q+1$. Moreover, any two distinct one-dimensional subspaces generate a unique two dimensional subspace so that this gives us $2-(q^2+q+1, q+1, 1)$ design for any prime power q. This is called a *projective plane* of order q. This has a visual metaphor. A projective plane of order n is a collection X of n^2+n+1 elements called "points" and a collection B of n^2+n+1 blocks called "lines". We require that each point is on precisely $n+1$ lines and each line has precisely $n+1$ points, and any two distinct points determine a unique line. Thus, a projective plane of order n is a $2-(n^2+n+1, n+1, 1)$ design. It is unknown if there are any projective planes of order n when n is not a prime power. We will address this question below using the Bruck-Ryser-Chowla theorem.

The *Fano plane* consisting of seven points and seven lines is the $2-(7,3,1)$ design constructed above using the finite field of two elements. This is usually represented by a triangle along with the midpoints of

the three sides together with the centroid. The lines are the sides of the triangle, the lines joining the midpoints of the sides and finally the "line" joining the three midpoints usually drawn as a circle. This has the amusing application to the following problem. Arrange the luncheon engagements of seven people over a week long period in such a way that each day three people have lunch together and by the end of the week, any two of the people would have had lunch together precisely once. If we think of the Fano plane and view the vertices as the people, the lines representing the days of the week, the points on the line determine which of the three people should lunch together, then we have a visual resolution of the required arrangement.

We now prove the only non-existence theorem known in the theory of projective planes.

THEOREM 9.4.1. *If a projective plane of order n exists and $n \equiv 1$ or $2 \pmod 4$, then n can be expressed as a sum of two squares.*

PROOF. As observed earlier, we are asking for the existence of a $2-(n^2+n+1, n+1, 1)$ design. Notice that $v = n(n+1)+1$ is odd. Applying the Bruck-Ryser-Chowla theorem, we deduce that the Diophantine equation

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

has a non-trivial integral solution. If $n \equiv 1 \pmod 4$, then $n(n+1)/2$ is odd so the theorem says that we can solve

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

in non-zero integers. The same implication occurs when $n \equiv 2 \pmod 4$. Thus n is the sum of two rational squares. To complete the proof, we need to show that n is in fact the sum of two integral squares. Now we need to use one more fact from number theory. Recall that an odd prime number p can be written as a sum of two squares if and only if $p \equiv 1 \pmod 4$. From this, one can deduce that the numbers that can be expressed as a sum of two integer squares are precisely the numbers whose unique factorization into distinct prime powers does not admit a prime $\equiv 3 \pmod 4$ to an odd power. Thus, if n cannot be written as a sum of two squares, then there is a prime $p \equiv 3 \pmod 4$ an odd power p^{2a+1} (say) of which divides n exactly. Reducing the equation mod p^{2a+1}, we get

$$y^2 + z^2 \equiv 0 \pmod{p^{2a+1}}.$$

If y, z are coprime to p, this is already a contradiction for it says that -1 is a perfect square mod p. If y and z are not coprime to p, only an

even power of p can divide each of them and hence both of them and after canceling it, we get a contradiction that completes the proof. ∎

We can apply this result to show that there is no projective plane of order 6. Indeed, if there is, by the previous theorem, 6 can be written as a sum of two integral squares, which is clearly not the case. Thus, there is no $2-(43,7,1)$ design. In particular, there is no way to arrange 43 objects into 43 blocks such that each block contains 7 objects and any two objects occurring together in precisely one block.

For a long time, the first unresolved case was $n = 10$. The above theorem does not exclude this possibility as 10 can be written as $1+9$. In 1991, Clement Lam of Concordia University, Canada using the Cray 1 computer showed that there is no projective plane of order 10. Thus, we still have no conceptual proof of this fact. It is generally believed that projective planes can only exist when n is a prime power, but this has not yet been proved.

9.5. Proof of the Bruck-Ryser-Chowla Theorem

The proof of Theorem 9.3.4 requires the use of Lagrange's four square theorem. This theorem says that every natural number can be written as a sum of four squares of natural numbers. We prove it in four steps. As the identity

$$(|z|^2+|w|^2)(|u|^2+|w|^2) = |uz-\overline{w}v|^2+|wu+\overline{z}v|^2$$

is easy to verify directly for all complex numbers u, v, w, z, we deduce from it, by putting $z = x_1+ix_2$, $w = x_3+ix_4$, $u = y_1+iy_2$, $w = y_3+iy_4$ 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

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

This means that if a can be written as a sum of four integral squares, and b can be written as a sum of four integral squares, so can ab and we have an explicit recipe for determining these squares if we know the ones for a and b respectively. As every number is a product of prime numbers, it therefore suffices to prove Lagrange's theorem for prime numbers.

The next step is to see that for any odd prime p, we can solve the congruence

$$x^2 + y^2 + 1 \equiv 0 \pmod{p}.$$

To see this, we consider the set of squares mod p, which has size $1+(p-1)/2 = (p+1)/2$. The same is true of the set of elements of the form $-1-y^2$. If these sets were disjoint, we would get at least $p+1$ residue classes mod p, a contradiction. Hence, there is a common element and this gives a solution to the congruence. Since the integers in the interval $[-(p-1)/2, (p-1)/2]$ forms a complete set of residue classes mod p, we may choose $|x| < p/2$ and $|y| < p/2$, we deduce that there are integers x, y so that

$$x^2 + y^2 + 1 = mp$$

with $m < p$.

The third step is to consider the smallest natural number m such that mp can be written as a sum of four squares. By the previous paragraph, the set is non-empty. Call the smallest such m, m_0. Then, $m_0 < p$. If $m_0 = 1$, we are done so let us suppose that $1 < m_0 < p$. Hence, we can write

$$m_0 p = x_1^2 + x_2^2 + x_3^2 + x_4^2.$$

If m_0 were even, then either all of the x_i's are even or all of them are odd, or precisely two of them, say, x_1, x_2 (without loss of generality) are even. In any of the cases, $x_1 - x_2, x_1 + x_2, x_3 - x_4, x_3 + x_4$ are even and we have

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

Thus $(m_0/2)p$ can be written as a sum of four squares and this is a contradiction to the minimality of m_0. So we may suppose m_0 is odd.

The final step involves choosing y_1, y_2, y_3, y_4 so that $y_i \equiv x_i \pmod{m_0}$ with $|y_i| \leq (m_0 - 1)/2$. Then,

$$m_0 m_1 = y_1^2 + y_2^2 + y_3^2 + y_4^2$$

with $m_1 < m_0$. By step 1, we see that $(m_0 p)(m_0 m_1)$ can be written as a sum of four squares:

$$z_1^2 + z_2^2 + z_3^2 + z_4^2$$

with the z_i's being given explicitly in terms of x_i's and the y_i's. From this explicit description, we see directly that $z_i \equiv 0 \pmod{m_0}$. Thus, we may divide out by m_0^2 and deduce that $m_1 p$ can be written as a sum of four squares. But this contradicts the minimality of m_0 as $m_1 < m_0$. This completes the proof of Lagrange's theorem.

Now we will sketch the proof of the Bruck-Ryser-Chowla theorem. Suppose that we have a symmetric (v, k, λ) design with v odd. Let $n = k - \lambda$ and suppose that $v \equiv 3 \pmod 4$. We want to show that

$$nx^2 = z^2 + \lambda y^2$$

has a non-trivial integral solution. It suffices to show that this has a non-trivial rational solution, since we can always clear denominators.

By Lagrange's theorem, we may write $n = a^2 + b^2 + c^2 + d^2$ and so let H be the 4×4 matrix:

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

Then, $HH^t = H^tH = nI$. Now let A be the incidence matrix of the symmetric block design. This is a $v \times v$ matrix. Now look at the $(v+1) \times (v+1)$ matrix B obtained by adding a 1 in the $(v+1, v+1)$-th position and zeros everywhere else in the last row and last column. Then,

$$B^tB = \begin{pmatrix} A^tA & 0 \\ 0 & 1 \end{pmatrix}.$$

As $4|v+1$, we may create the $(v+1) \times (v+1)$ matrix K which has $(v+1)/4$ diagonal blocks of the matrix H. Then, $K^tK = KK^t = nI$. Consider the quadratic form

$$\mathrm{x}^tB^tB\mathrm{x} = k(x_1^2 + \cdots + x_v^2) + x_{v+1}^2 + \lambda \sum_{i \neq j \leq v} x_ix_j.$$

If we put $\mathrm{z} = B\mathrm{x}$, then this is

$$\sum_i z_i^2.$$

We may "complete squares" and re-write this as

$$\lambda(x_1 + \cdots + x_v)^2 + x_{v+1}^2 + n(x_1^2 + \cdots + x_v^2).$$

Consider another change of co-ordinates: $\mathrm{z} = K\mathrm{y}$. Then

$$\mathrm{z}^t\mathrm{z} = \mathrm{y}^tK^tK\mathrm{y}$$

which is

$$\sum_i z_i^2 = n \sum_i y_i^2.$$

Thus, $\mathrm{x} = (B^{-1}K)\mathrm{y}$ so that

$$n(y_1^2 + \cdots y_{v+1}^2) = \lambda(x_1 + \cdots + x_v)^2 + x_{v+1}^2 + n(x_1^2 + \cdots + x_v^2).$$

The idea now is to choose the x_i and y_i suitably so as to obtain the statement of the theorem. As the matrix $B^{-1}K$ is a rational matrix, we may write

$$x_i = \sum_{i \geq 1} a_i y_i$$

with a_i rational. If $a_1 \neq 1$, choose $x_1 = y_1$; otherwise, choose $x_1 = -y_1$. In either case, $x_1^2 = y_1^2$ and y_1 is a rational linear combination of $y_2, ..., y_{v+1}$. Thus, x_2 is a rational linear combination of $y_2, ..., y_{v+1}$:

$$x_2 = \sum_{i \geq 2} b_i y_i$$

with b_i rational. If $b_2 \neq 1$, choose $x_2 = y_2$; otherwise, choose $x_2 = -y_2$. In either case $x_2^2 = y_2^2$ and y_2 is now a rational linear combination of $y_3, ..., y_{v+1}$. We continue in this way for each $i \leq v$ so that $x_i^2 = y_i^2$ for each $i \leq v$ and y_v is a rational multiple of y_{v+1} and x_{v+1} is a rational multiple of y_{v+1}. Put $y_{v+1} = 1$. Then, x_{v+1} and y_v are uniquely determined rational numbers and working backwards, so are all the x_i's and the y_i's. Since $x_i^2 = y_i^2$ for $1 \leq i \leq v$, we get

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

has a solution in rational numbers. Moreover, the solution is non-trivial since x_{v+1} and y_{v+1} are non-zero. This completes the proof in this case.

The case $v \equiv 1 \pmod 4$ is similar and we leave it as an exercise to the reader. The essential change in the above proof is that we use the matrix A instead of the matrix B and replace K by the $v \times v$ matrix obtained by putting H on the diagonal and adding a 1 in the (v, v) position and zeros elsewhere in the last row and column. Then, the proof proceeds as before and we leave it as an exercise to the reader.

9.6. Codes and Designs

The fundamental paper *A mathematical theory of communications* from 1948 of Claude Shannon (1916-2001) is considered to be the starting point of coding theory. Around the same time, Richard Wesley Hamming (1915-1998) and Marcel J.E. Golay (1902-1989) also contributed to the beginning of this subject.

A **code** is a subset of $\mathbf{F}_q^n$. A code is called **linear** if it is a subspace of $\mathbf{F}_q^n$. It is **binary** if $q = 2$. The vectors in the code are called **codewords**. The **weight** of a vector v, denoted $wt(v)$, is the number of non-zero coordinates of v. The **Hamming distance** between two vectors v and w is the weight of $v - w$, and is denoted $d(v, w)$. If C is a code, the minimum distance $d(C)$ is the minimum of $d(v, w)$ for v, w distinct elements of C.

A code is said to be e-**error correcting** if $d(C) \geq 2e+1$. The reason for this definition is given by the following theorem.

THEOREM 9.6.1. *A code is e-error correcting if and only if the Hamming spheres:*

$$B_e(c) := \{v : d(v,c) \leq e\}$$

are disjoint for all $c \in C$.

PROOF. If $B_e(c_1)$ and $B_e(c_2)$ are not disjoint for two distinct codes c_1, c_2, then let v be a common element of these two Hamming spheres. Then,

$$d(c_1,c_2) \leq d(c_1,v) + d(v,c_2) \leq 2e.$$

But $d(c_1,c_2) \geq 2e+1$ for any two distinct code words, so this is a contradiction.

Conversely, if all the Hamming spheres are disjoint, and C is not e-error correcting, then there are two codewords c_1, c_2 such that $d(c_1,c_2) = f \leq 2e$. This means that c_1 and c_2 do not agree in f positions. Now change the co-ordinates of c_1 in $\lfloor f/2 \rfloor$ of these positions to agree with c_2 and call this changed vector b. Because $f \leq 2e$, we have that

$$d(c_1,b) = \lfloor f/2 \rfloor \leq e, \quad d(c_2,b) = f - \lfloor f/2 \rfloor \leq e$$

so that b is an element of $B_e(c_1)$ and $B_e(c_2)$ which is a contradiction. ∎

The application of these ideas in communication networks is as follows. If C is an e-error correcting code, then these codewords are used to send signals over a "noisy channel". If a code word c is received as c' and e errors are made in the transmission, then $d(c,c') \leq e$. Thus c' lies in the Hamming sphere $B_e(c)$. By Theorem 9.6.1, this is the unique code word satisfying this inequality.

We can construct error correcting codes by taking the rows of the incidence matrix of a symmetric (v,k,λ)-design as code words. Any two words have λ 1's together in precisely λ places. Each code has precisely k 1's and $v-k$ 0's. If R_1 and R_2 are distinct rows, then the number of co-ordinates with entry 1 at which R_1 and R_2 agree is the dot product $R_1 \cdot R_2$ and this is λ. If J is the vector consisting of all 1's, then the number of co-ordinates with entry 0 at which R_1 and R_2 agree is the dot product $(J-R_1)\cdot(J-R_2)$ which is $v-2k+\lambda$. By the definition of the Hamming distance, we deduce that

$$d(R_1,R_2) = 2(k-\lambda).$$

Thus, the rows of a symmetric (v,k,λ) design give us a $(k-\lambda-1)$-error correcting code.

In 1971, the Mars Mariner spacecraft used the rows of a $(31, 15, 7)$ design as codewords to send back photographs of Mars back to Earth. This code corrects 7 errors. In later space missions, more sophisticated codes called Reed-Solomon codes have been used and these codes are capable of correcting a larger number of errors. They are based on the following simple idea. Given a code word $(a_0, a_1, ..., a_{m-1})$, construct a polynomial

$$f(x) = a_0 + a_1 x + \cdots + a_{m-1} x^{m-1}.$$

Fix a primitive root g of $\mathbf{F}_q$. Instead of trying to send the code word, the spacecraft transmits the sequence $f(0), f(g), \ldots, f(g^N)$ where $N > m$. Since a polynomial of degree m is determined by $m + 1$ values, this is sufficient information to retrieve the original code word $(a_0, ..., a_{m-1})$ and this can be done algorithmically in an efficient way. One can prove that this method gives rise to a $(q + m)/2$-error correcting code.

9.7. Exercises

EXERCISE 9.7.1. Prove that

$$(q^k - 1)\binom{n}{k}_q = (q^n - 1)\binom{n-1}{k-1}_q.$$

EXERCISE 9.7.2. Prove that

$$\binom{n+1}{k}_q = \binom{n}{k-1}_q + \binom{n}{k}_q + (q^n - 1)\binom{n-1}{k-1}_q.$$

EXERCISE 9.7.3. Let $f_q(n)$ be the number of subspaces of $\mathbb{F}_q^n$. Show that

$$f_q(n+1) = 2f_q(n) + (q^n - 1)f_q(n-1).$$

EXERCISE 9.7.4. Let L be the lattice of subspaces of $\mathbb{F}_q^n$ partially ordered by inclusion. If W is a subspace of dimension k, show that

$$\mu(0, W) = (-1)^k q^{\binom{k}{2}}.$$

EXERCISE 9.7.5. 16 students decide to sign up for three fields each. Each trip accommodates precisely 6 students. The students would like to sign up in such a way that any two of them would be together on precisely one of the trips. Is such arrangement possible ? Explain.

EXERCISE 9.7.6. Construct explicitly a $2-(31, 3, 1)$ design. For any natural number $n \geq 1$, show that there exists a $2-(2^n - 1, 3, 1)$ design.

EXERCISE 9.7.7. If A is a $v \times b$ matrix and B is a $b \times v$ matrix, show that

$$\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B)).$$

EXERCISE 9.7.8. In a symmetric $2-(v,k,\lambda)$ design with incidence matrix A, show that

$$\frac{1}{k-\lambda}\left(A+\sqrt{\frac{\lambda}{k}}J\right)$$

is the inverse of

$$A^t-\sqrt{\frac{\lambda}{v}}J.$$

Deduce that

$$A^tA=\lambda J+(r-\lambda)I.$$

Use this equation to prove that in any symmetric design, every pair of blocks has precisely λ elements in common.

EXERCISE 9.7.9. Show that there is no projective plane of order 14.

EXERCISE 9.7.10. If $p\equiv 3 \pmod 4$ is a prime, show that there is no $2-(v,p+1,1)$ design with $v\equiv 3 \pmod 4$.

EXERCISE 9.7.11. If C is a code in $\mathbb{F}_q^n$ with distance $d(C)\geq 2e+1$, then

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

EXERCISE 9.7.12. If C is a code in $\mathbb{F}_q^n$ with distance $d(C)=d$, then

$$|C|\leq q^{n-d+1}.$$

EXERCISE 9.7.13. Label the points of the Fano plane by the elements of $\mathbb{Z}_7$ such that each block of the Fano plane has the form $\{x,x+1,x+3\}$ for $x\in\mathbb{Z}_7$.

EXERCISE 9.7.14. Consider the following incidence structure: the points are the edges of the complete graph K_6 and the blocks are all the sets of three edges that form a perfect matching or a triangle in K_6. Show that this is a Steiner triple system on 15 points.

EXERCISE 9.7.15. Show that if $x,y,z\in\mathbb{F}_q^n$, then

$$d(x,z)\leq d(x,y)+d(y,z).$$

EXERCISE 9.7.16. Show that if a $2-(v,3,1)$ design exists, then $v\equiv 1,3 \pmod 6$.

EXERCISE 9.7.17. Consider the design whose point set is $\mathbb{Z}_n\times\mathbb{Z}_3$. The blocks are the triples

$$\{(x,0),(x,1),(x,2)\}$$

for $x \in \mathbb{Z}_n$ and

$$\left\{(x,i),(y,i),\left(\frac{x+y}{2},i+1\right)\right\}$$

for $x \neq y \in \mathbb{Z}_n$ and $i \in \mathbb{Z}_3$. Show that this is a $2-(6t+3,3,1)$-design.

EXERCISE 9.7.18. Show that the number of blocks in a $t-(v,k,\lambda)$ design is

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

EXERCISE 9.7.19. Show that in any $t-(v,k,1)$ design

$$v \geq (t+1)(k-t+1).$$

EXERCISE 9.7.20. Show that there are at most two disjoint Steiner triple systems on a set of 7 points.