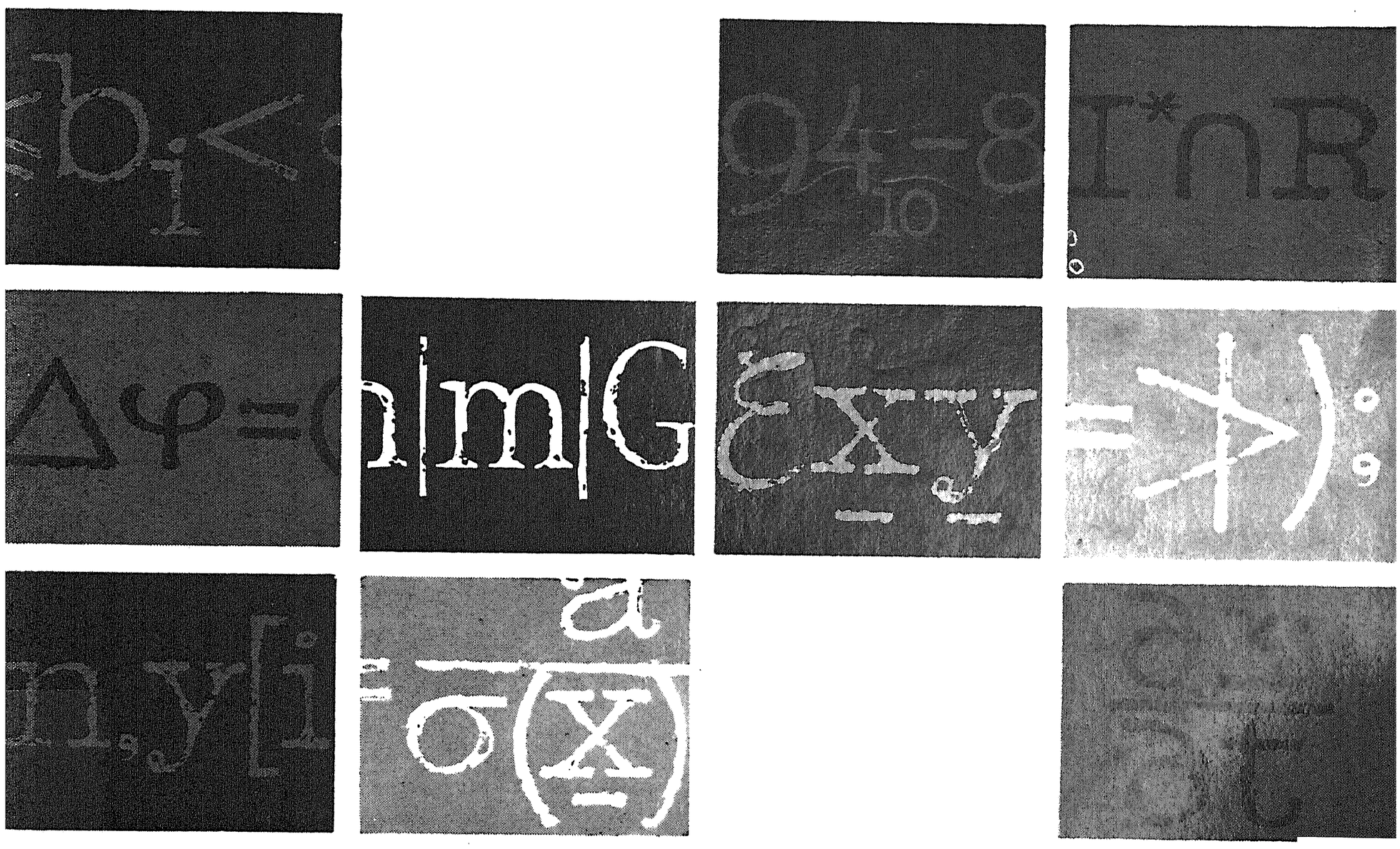
COMBINATORICS

Part 1: Theory of designs, finite geometry and coding theory

M. HALL, Jr. (ed.)
J.H. VAN LINT (ed.)



MATHEMATICAL CENTRE TRACTS 55

MATHEMATICAL CENTRE TRACTS 55

M. HALL, Jr. (ed.)
J.H. VAN LINT (ed.)

COMBINATORICS

Part 1: Theory of designs, finite geometry and coding theory

Proceedings of the Advanced Study Institute on Combinatorics held at Nijenrode Castle, Breukelen, The Netherlands, July 8-20, 1974

2nd edition (revised)

MATHEMATISCH CENTRUM AMSTERDAM 1975

AMS(MOS) subject classification scheme (1970): 05-02, 05B20, 05B05, 05B25, 94A10

1st edition: June 1974

2nd edition (revised): June 1975

ISBN 90 6196 099 1

CONTENTS

PREFACE

Combinatorics has come of age. It had its beginnings in a number of puzzles which have still not lost their charm. Among these are EULER's problem of the 36 officers and the KÖNIGSBERG bridge problem, BACHET's problem of the weights, and the Reverend T.P. KIRKMAN's problem of the schoolgirls. Many of the topics treated in ROUSE BALL's *Recreational Mathematics* belong to combinatorial theory.

All of this has now changed. The solution of the puzzles has led to a large and sophisticated theory with many complex ramifications. And it seems probable that the four color problem will only be solved in terms of as yet undiscovered deep results in graph theory. Combinatorics and the theory of numbers have much in common. In both theories there are many problems which are easy to state in terms understandable by the layman, but whose solution depends on complicated and abstruse methods. And there are now interconnections between these theories in terms of which each enriches the other.

Combinatorics includes a diversity of topics which do however have interrelations in superficially unexpected ways. The instructional lectures included in these proceedings have been divided into six major areas: 1. *Theory of designs*; 2. *Graph theory*; 3. *Combinatorial group theory*; 4. *Finite geometry*; 5. *Foundations, partitions and combinatorial geometry*; 6. *Coding theory*. They are designed to give an overview of the classical foundations of the subjects treated and also some indication of the present frontiers of research.

Without the generous support of the North Atlantic Treaty Organization, this *Advanced Study Institute on Combinatorics* would not have been possible, and we thank them sincerely. Thanks are also due to the National Science Foundation for the support of some advanced students, in addition to the support of those with their own NSF grants. The IBM Corporation has kindly given us financial support to supplement the NATO grant. The Xerox Corporation has helped with donations of material and equipment.

Finally we must acknowledge the extensive activities of the Mathematical Centre of Amsterdam in making all the arrangements necessary for holding this conference and preparing these proceedings.

M. HALL, Jr.

J.H. VAN LINT

THEORY OF DESIGNS

MATHEMATICAL CENTRE TRACTS 55, 1974, 3-17

INDETERMINATES AND INCIDENCE MATRICES *)

H.J. RYSER

California Institute of Technology, Pasadena, Cal. 91109, USA

1. INTRODUCTION

We let

(1.1) $$X = \{x_1,\ldots,x_n\}$$

denote a non-empty set of n elements. We call such a set an n-*set*. We let

(1.2) $$X_1,\ldots,X_m$$

denote m not necessarily distinct subsets of X. We refer to this collection of subsets of X as a *configuration* and remark that configurations occur in great profusion throughout the combinatorial literature.

We now let F denote an arbitrary field. We interconnect F and our configuration by regarding the elements $x_1,\ldots,x_n$ of X as n independent indeterminates with respect to the field F. This simple device immediately imposes an algebraic structure on our original configuration and allows us to carry out various algebraic manipulations within the polynomial ring

(1.3) $$F[x_1,\ldots,x_n].$$

We exploit this algebraic structure further by the introduction of an incidence matrix A. We set $a_{ij} = 1$ if $x_j \in X_i$ and we set $a_{ij} = 0$ if $x_j \notin X_i$. In these equations the 1 and the 0 are the identity element and the zero element, respectively, of the field F. The resulting matrix

*) This research was supported in part by the Army Research Office-Durham under Grant DA-ARO-D-31-124-72-G171 and the National Science Foundation under Grant GP-36230X.

(1.4) $$A = [a_{ij}] \qquad (i=1,\ldots,m;\ j=1,\ldots,n)$$

of size $m \times n$ is the *incidence matrix* for the subsets $X_1,\ldots,X_m$ of X. Row i of A displays the subset X_i and column j of A displays the occurrences of the element x_j among the subsets. Thus A gives us a complete description of the subsets and the occurrences of the elements within the subsets.

We may now write

(1.5) $$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} ,$$

where the component y_i of the vector on the left side of (1.5) is precisely the sum of the elements of the subset X_i of X. The equation (1.5) opens the door for important matrix manipulations. For example, equation (1.5) and its transpose allow us to associate with our configuration the quadratic form

(1.6) $$y_1^2 + \ldots + y_m^2 = (x_1,\ldots,x_n)A^TA \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} .$$

We point out that this quadratic form has been exploited with great success in the study of block designs. (See, for example, [3], [11], [19].)

Indeterminates may be associated with a given configuration in various ways. In what follows we discuss in some detail two such associations. The one deals with n independent indeterminates that represent the elements of the n-set X and the other involves mn independent indeterminates associated with the positions of the incidence matrix A of size $m \times n$. Much of the material that we discuss is in the very early stages of its development. But we anticipate that these topics hold great potential for further study. Our concluding remarks deal with indeterminates and the incidence matrix A of a (v,k,λ)-design.

2. A FUNDAMENTAL MATRIX EQUATION FOR FINITE SETS

We return to a configuration of subsets $X_1,\ldots,X_m$ of an n-set $X = \{x_1,\ldots,x_n\}$, where we regard the elements $x_1,\ldots,x_n$ as n independent indeterminates with respect to a field F. This configuration now has an

incidence matrix A of size m × n. We also denote by X the diagonal matrix of order n

$$X = \mathrm{diag}[x_1,\ldots,x_n], \tag{2.1}$$

and we then form the matrix equation

$$AXA^T = Y, \tag{2.2}$$

where A^T denotes the transpose of the matrix A. All of the matrices in (2.2) have elements in the polynomial ring $F[x_1,\ldots,x_n]$.

The fundamental matrix equation (2.2) is a most remarkable one because it contains a vast amount of information in a highly compact form. The matrix Y is a symmetric matrix of order m and we know the structure of this matrix explicitly. Thus the matrix Y has in its (i,j) position the sum of the indeterminates in the set intersection $X_i \cap X_j$, and it follows that the matrix Y gives us an explicit representation for all of these set intersections. In particular, the elements on the main diagonal of Y display the subsets $X_1,\ldots,X_m$ of our original configuration. The matrix equation (2.2) was introduced by RYSER in [20] and [22]. Some other investigations that deal with matrices and set intersections include [9], [10], [14], [21].

The matrix Y involves the indeterminates $x_1,\ldots,x_n$ and we write

$$Y = Y(x_1,\ldots,x_n). \tag{2.3}$$

We may assign the indeterminates in the matrix equation (2.2) arbitrary values of the field F, and each such assignment produces a new matrix equation that must be satisfied by the incidence matrix A of our configuration. Suppose, for example, that we have m = n. Then A is a square matrix of order n and suppose further that the matrix A is non-singular. Then if we assign x_i the value e_i in F it follows that the matrix $Y(e_1,\ldots,e_n)$ is congruent to the diagonal matrix

$$E = \mathrm{diag}[e_1,\ldots,e_n] \tag{2.4}$$

with respect to the field F, and this congruence relationship remains valid for arbitrary choices of the e_i in F.

In many problems it is desirable to select F as the field of rational numbers or some extension field of this field. The incidence matrix A is then a (0,1)-*matrix* of size m × n. If we now set $x_1 = \ldots = x_n = 1$, then

(2.2) reduces to the classical equation

$$AA^T = Y(1,\ldots,1). \tag{2.5}$$

In this case the matrix $Y(1,\ldots,1)$ on the right-hand side of (2.5) reveals the cardinalities of the set intersections.

The following theorem appears in [20] and affords a good illustration of the type of result that is motivated by the matrix equation (2.2). The proof of the theorem uses techniques similar to those employed by VAN LINT & RYSER [15] in their study of block designs with repeated blocks.

THEOREM 2.1. *Suppose that* A *is a* (0,1)*-matrix of order* n *and suppose that* A *satisfies the matrix equation*

$$AEA^T = D, \tag{2.6}$$

where A^T *is the transpose of the matrix* A. *Suppose further that the matrices* D *and* E *are real (or complex) diagonal matrices of order* n *and that* D *is non-singular. Then it follows that* A *is a permutation matrix.*

PROOF. The matrix D is non-singular so that we may write

$$AEA^TD^{-1} = EA^TD^{-1}A = I, \tag{2.7}$$

where I is the identity matrix of order n. Hence it follows that

$$A^TD^{-1}A = E^{-1}. \tag{2.8}$$

We now let

$$D = \mathrm{diag}[d_1,\ldots,d_n], \quad E = \mathrm{diag}[e_1,\ldots,e_n] \tag{2.9}$$

and inspect the main diagonal of (2.8). Then we obtain

$$A^T \begin{pmatrix} 1/d_1 \\ \vdots \\ 1/d_n \end{pmatrix} = \begin{pmatrix} 1/e_1 \\ \vdots \\ 1/e_n \end{pmatrix}. \tag{2.10}$$

We note that in (2.10) we have made strong use of the fact that A is a (0,1)-matrix. We now multiply (2.10) by AE and this gives

$$(2.11) \qquad A\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} .$$

Thus each of the row sums of the matrix A is equal to 1. But A is a non-singular (0,1)-matrix and hence it follows that A is a permutation matrix.□

Thus far our discussion has been motivated entirely by combinatorial considerations. But the matrix equation (2.2) also suggests some algebraic questions that are of considerable interest in their own right. The following theorem illustrates this point [22].

THEOREM 2.2. *Suppose that* Y *is a matrix of order* $n \geq 3$ *and such that every element of* Y *is a linear form in the* n *independent indeterminates* $x_1,\ldots,x_n$ *with respect to a field* F. *We let*

$$(2.12) \qquad X = \mathrm{diag}[x_1,\ldots,x_n]$$

and we suppose that the determinant of Y *satisfies*

$$(2.13) \qquad \det(Y) = cx_1 \cdots x_n,$$

where $c \neq 0$ *and* $c \in$ F. *We suppose further that every element of* Y^{-1} *is a linear form in* $x_1^{-1},\ldots,x_n^{-1}$ *with respect to the field* F. *Then there exist matrices* A *and* B *of order* n *with elements in* F *such that*

$$(2.14) \qquad AXB = Y.$$

The proof of theorem 2.2 is available in [22]. Further theorems that imply a factorization of Y of the form (2.14) are also valid. These results deal with compound matrices and do not require Y to be a square matrix [22]. Much earlier investigations by KANTOR [13], FROBENIUS [8], and SCHUR [24] study related problems but with X a matrix of size $m \times n$ and such that the elements of X are mn independent variables over the complex field. A more recent account of this older theory appears in [16].

3. THE FORMAL INCIDENCE MATRIX

We now let

$$(3.1) \qquad A = [a_{ij}] \qquad (i=1,\ldots,m;\ j=1,\ldots,n)$$

denote a matrix of size m × n with elements in a field F. We may still regard A as the *incidence matrix* for our configuration of m subsets of an n-set, where the non-zero elements of A play the role of the identity element of the field in our earlier representation. One of the most remarkable general theorems in combinatorial matrix theory is the following theorem of KÖNIG. (The theorem is also frequently referred to as the König-Egerváry theorem or the Frobenius-König theorem.) Throughout our discussion a *line* of a matrix designates either a row or a column of the matrix.

THEOREM 3.1. *Suppose that* A *is a matrix of size* m × n *with elements in a field* F. *Then the minimal number of lines in* A *that cover all of the non-zero elements in* A *is equal to the maximal number of non-zero elements in* A *with no two of the non-zero elements on a line.*

An enormous literature centers around this theorem and the related theorems of P. HALL [12], DILWORTH [4], and FORD & FULKERSON [7]. A detailed discussion of these topics is available in the recent book by MIRSKY [17]. KÖNIG's theorem is frequently stated in the terminology of (0,1)-matrices. But the nature of the theorem is such that it holds quite generally for an arbitrary rectangular array in which all of the elements of the array have been partitioned into exactly two components.

At this point we disregard our earlier notation and we let

$$X = [x_{ij}] \qquad (i=1,\ldots,m;\ j=1,\ldots,n) \tag{3.2}$$

denote the matrix of size m × n, where the elements of X are mn independent indeterminates with respect to the field F. We call the Hadamard product

$$M = A * X = [a_{ij}x_{ij}] \tag{3.3}$$

the *formal incidence matrix* associated with A. The elements of M belong to the polynomial ring

$$F^* = F[x_{11},x_{12},\ldots,x_{mn}]. \tag{3.4}$$

The formal incidence matrix is useful in various combinatorial investigations [2],[6],[18],[23],[26].

The maximal number of non-zero elements in A with no two of the non-zero elements on a line is called the *term rank* of A. It turns out that this important combinatorial invariant of A is equal to an algebraic invariant

of M. *The term rank of* A *is equal to the rank of* M. This observation is due to EDMONDS [6] and may be derived as follows. We note that a submatrix of M of order r has a non-zero determinant if and only if the corresponding submatrix of A has term rank r. This is a consequence of the definition of the formal incidence matrix. But the rank of a matrix is equal to the maximal order of a square submatrix with a non-zero determinant. Hence we obtain the desired conclusion.

We next discuss another basic combinatorial property of A in terms of an algebraic property of M. We now deal with square matrices of order n with elements in a field F. We say that a matrix A of order $n > 1$ is *fully indecomposable* provided that A does not contain a zero submatrix of size $r \times (n-r)$, for some integer r in the interval $1 \leq r \leq n-1$. It follows that the non-zero elements of a fully indecomposable matrix A of order $n > 1$ cannot be covered by n lines that are composed of both rows and columns of A. In case the matrix A is of order $n = 1$ then we say that A is *fully indecomposable* provided that A is not the zero matrix of order 1. We may conclude at once from theorem 3.1 that a fully indecomposable matrix A of order n has term rank n. Hence our earlier observation implies that a fully indecomposable matrix A has $\det(M) \neq 0$. But it is clear that $\det(M) \neq 0$ does not in general imply that the matrix A is fully indecomposable. However, the following theorem in a recent paper by RYSER [23] shows that a fully indecomposable matrix A is characterized by a somewhat deeper algebraic property of det(M).

THEOREM 3.2. *Suppose that* A *is a matrix of order* n *with elements in a field* F *and let* $M = A * X$ *denote the formal incidence matrix associated with* A. *Then the matrix* A *is fully indecomposable if and only if* det(M) *is an irreducible polynomial in the polynomial ring*

$$F^* = F[x_{11}, x_{12}, \ldots, x_{nn}]. \tag{3.5}$$

We do not attempt a derivation of theorem 3.2 here. But we remark that it is easy to show that if the polynomial det(M) is irreducable in F^* then the matrix A is fully indecomposable. The proof of the converse proposition is more difficult.

We describe briefly a lemma of some intrinsic interest used in the derivation of the converse. A *diagonal product* of a matrix of order n is a product of n elements of the matrix with no two of the elements on a line.

We now let X_r denote a submatrix of X of order r. We designate the u = r! diagonal products of X_r by

$$(3.6) \qquad y_1,\ldots,y_u.$$

We say that the polynomial f has an *indeterminate pattern* based on X_r provided that

$$(3.7) \qquad f = \sum_{i=1}^{u} a_i y_i,$$

where the coefficients a_i are in F and not all of the a_i are zero. A polynomial with an indeterminate pattern based on X_r is homogeneous and of degree r over F. We note that if det(M) $\neq$ 0, then det(M) is an example of a polynomial with an indeterminate pattern based on X. Two submatrices B and C of orders r and n-r, respectively, of a matrix A of order n are called *complementary* provided that they are formed from complementary sets of lines of A. We are now ready to state the lemma used in the derivation of theorem 3.2 [23].

LEMMA 3.1. *Suppose that* h *is a polynomial with an indeterminate pattern based on* X *and suppose that in* F^* *we have*

$$(3.8) \qquad h = fg,$$

where f *and* g *are polynomials of positive degrees* r *and* n-r, *respectively. Then it follows that the polynomials* f *and* g *have indeterminate patterns based on* X_r *and* X_{n-r}, *respectively, where* X_r *and* X_{n-r} *are complementary submatrices of* X *of orders* r *and* n-r, *respectively.*

We now let A denote a matrix of order n with elements in a field F and we suppose that A is of term rank n. Then it follows that there exist permutation matrices P and Q of order n such that

$$(3.9) \qquad PAQ = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ * & A_2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ * & * & \ldots & A_r \end{bmatrix},$$

where the matrices $A_1, A_2, \ldots, A_r$ are fully indecomposable. These matrices are

called the *fully indecomposable components* of A. A theorem of DULMAGE & MENDELSOHN [1],[5] asserts the following.

THEOREM 3.3. *Suppose that* A *is a matrix of order* n *with elements in a field* F *and suppose that* A *is of term rank* n. *Then the fully indecomposable components of* A *are unique apart from order and row and column permutations within components.*

PROOF. Our proof follows [23] and is based on algebraic properties of the formal incidence matrix M = A * X. The matrix A is of term rank n so that we know that det(M) $\neq$ 0. We let $A_1,\ldots,A_r$ and $B_1,\ldots,B_s$ denote two sets of fully indecomposable components of A. Suppose that we apply certain permutations to the rows and the columns of A and also apply the identical permutations to the rows and the columns of M. Then we observe that the zero elements in both of the permuted matrices occupy the identical positions. Thus we may write

$$\det(M) = \pm f_1 \cdots f_r = \pm g_1 \cdots g_s, \tag{3.10}$$

where each of the polynomials f_i and g_j in (3.10) has an indeterminate pattern based on the appropriate submatrix of X. Each of these polynomials uniquely describes its associated submatrix of X. Moreover, this submatrix of X corresponds in A to a fully indecomposable component of A. Hence by theorem 3.2 we may conclude that the polynomials f_i and g_j are irreducible polynomials in F^*. But F^* is a unique factorization domain and this means that r = s and the f_i and the g_j are the same apart from order and scalar factors. Hence it follows that the fully indecomposable components A_i and B_j are the same apart from order and row and column permutations within components. □

The preceding proof is especially intriguing because the uniqueness of the fully indecomposable components is now a natural consequence of the unique factorization property of the polynomial ring F^*.

Suppose that A is a fully indecomposable matrix of order n with elements in a field F and let M = A * X denote the formal incidence matrix associated with A. Then by theorem 3.2 we may associate with A the irreducible polynomial det(M) in F^*. This correspondence between fully indecomposable matrices and irreducible polynomials is a most remarkable one because the polynomial det(M) determines the matrix A uniquely apart from multiplic-

ation of A on the left and the right by diagonal matrices [23]. Our derivation of this fact requires the following theorem of SINKHORN & KNOPP [25].

THEOREM 3.4. *Suppose that* A *is a fully indecomposable matrix of order* n *with elements in a field* F *and suppose that all of the non-zero diagonal products of* A *are equal. Then there exists a unique matrix* B *of order* n *with non-zero elements and of rank one such that* $b_{ij} = a_{ij}$ *whenever* $a_{ij} \neq 0$.

The following theorem concerning the correspondence between fully indecomposable matrices and irreducible polynomials is now a fairly easy consequence of theorem 3.4. Details of the proof are available in [23].

THEOREM 3.5. *Suppose that* A *and* B *are fully indecomposable matrices of order* n *with elements in a field* F *and let* M = A * X *and* N = B * X *denote the formal incidence matrices associated with* A *and* B, *respectively. Suppose further that*

$$\det(M) = c \det(N) \neq 0, \tag{3.11}$$

where c *is a scalar in* F. *Then there exist diagonal matrices* D *and* E *with elements in* F *such that*

$$DAE = B. \tag{3.12}$$

4. SYMMETRIC BLOCK DESIGNS

We recall that a (v,k,λ)-*design (symmetric block design)* is a configuration of subsets $X_1,\ldots,X_v$ of a v-set $X = \{x_1,\ldots,x_v\}$ subject to the following postulates:

(4.1) each X_i is a k-subset of X;

(4.2) each $X_i \cap X_j$ for $i \neq j$ is a λ-subset of X;

(4.3) the integers v, k and λ satisfy $0 < \lambda < k < v-1$.

These postulates imply that the incidence matrix A of a (v,k,λ)-design is a (0,1)-matrix of order v that satisfies the matrix equation

$$AA^T = (k-\lambda)I + \lambda J, \tag{4.4}$$

where A^T is the transpose of the matrix A, I is the identity matrix of order v, and J is the matrix of 1's of order v. One may show that the

incidence matrix A of a (v,k,λ)-design is normal, namely,

$$AA^T = A^TA, \tag{4.5}$$

and that the parameters v, k and λ satisfy the relationship

$$k - \lambda = k^2 - \lambda v. \tag{4.6}$$

Of special importance are the (v,k,λ)-designs with $\lambda = 1$. These configurations are called *finite projective planes*. One of the main unresolved problems in the study of block designs is the determination of the precise range of values of v, k and λ for which (v,k,λ)-designs exist. Detailed discussions of these topics are available in the books by DEMBOWSKI [3], HALL [11], and RYSER [19].

In what follows we make some observations concerning indeterminates and the incidence matrix A of a (v,k,λ)-design. We look mainly at the matrix equation (2.2). Then we have

$$AXA^T = Y, \tag{4.7}$$

where

$$X = \mathrm{diag}[x_1,\ldots,x_v]. \tag{4.8}$$

An appropriate analysis of the matrix equation (4.7) could conceivably yield important breakthroughs concerning the non-existence of (v,k,λ)-designs. One possible attack is an ingenious assignment of values e_i to the indeterminates x_i so that the resulting matrix equation contains a contradiction.

We now prove a new result that illustrates this idea.

THEOREM 4.1. *Suppose that* A *is the incidence matrix of a* (v,k,λ)*-design and suppose that* A *satisfies the matrix equation*

$$AEA^T = C, \tag{4.9}$$

where E *is a diagonal matrix such that all of the diagonal elements of* E *are equal to* ± 1. *Suppose further that all of the off-diagonal elements of* C *are also equal to* ± 1. *Then it follows that* $\lambda = 1$ *and thus* A *is the incidence matrix of a finite projective plane.*

PROOF. The matrix A is the incidence matrix of a (v,k,λ)-design and A satisfies the matrix equation (4.9). We let c_i denote the element in

position i of the main diagonal of C and we inspect the main diagonal of (4.9). Then we obtain

$$(4.10) \qquad AE\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix} = \begin{pmatrix}c_1\\ \vdots\\ c_v\end{pmatrix}.$$

We note that in (4.10) we have again made strong use of the fact that A is a (0,1)-matrix. We now multiply (4.10) by its transpose and we thereby obtain

$$(4.11) \qquad AEJE^TA^T = [c_ic_j].$$

We next square the matrix equation (4.9) and by (4.4) and (4.5) we have

$$(4.12) \qquad AE((k-\lambda)I + \lambda J)EA^T = C^2.$$

But

$$(4.13) \qquad E^2 = I$$

and hence by (4.4) and (4.11) we have

$$(4.14) \qquad (k-\lambda)^2I + \lambda(k-\lambda)J + \lambda[c_ic_j] = C^2.$$

The off-diagonal elements of C are equal to ±1 and hence an inspection of the main diagonal of (4.14) implies

$$(4.15) \qquad (k-\lambda)^2 + \lambda(k-\lambda) + \lambda c_i^2 = c_i^2 + (v-1).$$

Then by (4.6) we have

$$(4.16) \qquad (\lambda-1)c_i^2 = \frac{k(k-1)}{\lambda} - k(k-\lambda) \geq 0.$$

We may rewrite (4.16) in the form

$$(4.17) \qquad \frac{\lambda^2-1}{\lambda} \geq \frac{k(\lambda-1)}{\lambda}.$$

Now the assumption $\lambda > 1$ implies

$$(4.18) \qquad \lambda \geq k-1.$$

But by (4.3) we have $\lambda \leq k-1$ and hence $\lambda = k-1$. But then $k = v-1$ and this contradicts (4.3). Hence we have $\lambda = 1$ and A is the incidence matrix of a finite projective plane. □

We note that a converse type proposition is immediate. Thus suppose that A is the incidence matrix of a finite projective plane. Then each off-diagonal element of the matrix Y in (4.7) is equal to some indeterminate x_i. Hence it follows that an arbitrary assignment of values ±1 to the indeterminates x_i in (4.7) gives us a matrix equation of the form (4.9) that satisfies all of the requirements of theorem 4.1.

We conclude with a few remarks on the formal incidence matrix

$$M = A * X = [a_{ij}x_{ij}], \tag{4.19}$$

where A is the incidence matrix of a (v,k,λ)-design and

$$X = [x_{ij}] \tag{4.20}$$

is the matrix of v^2 independent indeterminates with respect to the rational field. It is easy to verify that the incidence matrix A of a (v,k,λ)-design is fully indecomposable and hence by theorem 3.2 we have that det(M) is an irreducible polynomial in the polynomial ring of the v^2 independent indeterminates with respect to the rational field.

One is now tempted to study the matrix equations

$$(A * X)A^T = Y(x_{11},x_{12},\ldots,x_{vv}) \tag{4.21}$$

and

$$(A * X)(A * X)^T = Z(x_{11},x_{12},\ldots,x_{vv}). \tag{4.22}$$

These matrix equations afford vastly greater substitution possibilities than does the matrix equation (4.7). For example, suppose that A is the incidence matrix of a finite projective plane and that we replace the matrix X on the left sides of (4.21) and (4.22) by an arbitrary (1,-1)-matrix of order v. Then the resulting matrices on the right sides of (4.21) and (4.22) have the property that all of their off-diagonal elements are also equal to ±1.

REFERENCES

[1] BRUALDI, R.A., *Permanent of the product of doubly stochastic matrices*, Proc. Cambridge Philos. Soc., 62 (1966) 643-648.

[2] BRUALDI, R.A. & H. PERFECT, *Extension of partial diagonals of matrices I*, Monatsh. Math., 75 (1971) 385-397.

[3] DEMBOWSKI, P., *Finite geometries*, Ergebnisse der Mathematik 44, Springer Verlag, Berlin, 1968.

[4] DILWORTH, R.P., *A decomposition theorem for partially ordered sets*, Ann. of Math. (2), 51 (1950) 161-166.

[5] DULMAGE, A.L. & N.S. MENDELSOHN, *Coverings of bipartite graphs*, Canad. J. Math., 10 (1958) 517-534.

[6] EDMONDS, J., *Systems of distinct representatives and linear algebra*, J. Res. Nat. Bur. Standards Ser. B, 71 (1967) 241-245.

[7] FORD Jr., L.R. & D.R. FULKERSON, *Flows in networks*, Princeton University Press, Princeton, 1962.

[8] FROBENIUS, G., *Über die Darstellung der endlichen Gruppen durch lineare Substitutionen*, Sitzungsberichte Berliner Akademie (1897) 994-1015.

[9] GOODMAN, A.W., *Set equations*, Amer. Math. Monthly, 72 (1965) 607-613.

[10] HALL Jr., M., *A problem in partitions*, Bull. Amer. Math. Soc., 47 (1941) 804-807.

[11] HALL Jr., M., *Combinatorial theory*, Blaisdell, Waltham, Mass., 1967.

[12] HALL, P., *On representatives of subsets*, J. London Math. Soc., 10 (1935) 26-30.

[13] KANTOR, S., *Theorie der Äquivalenz von linearen ∞^{λ}-Scharen bilinearer Formen*, Sitzungsberichte Münchener Akademie (1897) 367-381.

[14] KELLY, J.B., *Products of zero-one matrices*, Canad. J. Math., 20 (1968) 298-329.

[15] LINT, J.H. VAN & H.J. RYSER, *Block designs with repeated blocks*, Discrete Math., 3 (1972) 381-396.

[16] MARCUS, M. & F. MAY, *On a theorem of I. Schur concerning matrix transformations*, Arch. Math. (Basel), 11 (1960) 401-404.

[17] MIRSKY, L., *Transversal theory*, Academic Press, New York, 1971.

[18] PERFECT, H., *Symmetrized form of P. Hall's theorem on distinct representatives*, Quart. J. Math. Oxford Ser. 2, 17 (1966) 303-306.

[19] RYSER, H.J., *Combinatorial mathematics*, Carus Math. Monograph No. 14, Math. Assoc. Amer., Wiley, New York, 1963.

[20] RYSER, H.J., *A fundamental matrix equation for finite sets*, Proc. Amer. Math. Soc., 34 (1972) 332-336.

[21] RYSER, H.J., *Intersection properties of finite sets*, J. Combinatorial Theory A, 14 (1973) 79-92.

[22] RYSER, H.J., *Analogs of a theorem of Schur on matrix transformations*, J. Algebra, 25 (1973) 176-184.

[23] RYSER, H.J., *Indeterminates and incidence matrices*, Linear and Multilinear Algebra, 1 (1973) 149-157.

[24] SCHUR, I., *Einige Bemerkungen zur Determinantentheorie*, Sitzungsberichte Berliner Akademie (1925) 454-463.

[25] SINKHORN, R. & P. KNOPP, *Problems involving diagonal products in nonnegative matrices*, Trans. Amer. Math. Soc., 136 (1969) 67-75.

[26] TUTTE, W.T., *The factorization of linear graphs*, J. London Math. Soc., 22 (1947) 107-111.

MATHEMATICAL CENTRE TRACTS 55, 1974, 18-41

CONSTRUCTIONS AND USES OF PAIRWISE BALANCED DESIGNS *)

R.M. WILSON

Ohio State University, Columbus, Ohio 43210, USA

1. INTRODUCTION

A *pairwise balanced design* (PBD) of index unity is a pair $(X,\mathcal{A})$ where X is a set (of *points*) and $\mathcal{A}$ a class of subsets A of X (called *blocks*) such that any pair of distinct points of X is contained in exactly one of the blocks of $\mathcal{A}$ (and we may also require $|A| \geq 2$ for each $A \in \mathcal{A}$). Such systems are also known as linear spaces. PBD's where all blocks have the same size $|A| = k$ are known as *balanced incomplete block designs* (BIBD's) of index $\lambda = 1$, as $2 - (v,k,1)$ designs, and as Steiner systems $S(2,k,v)$. The more general concept, where multiple block sizes are allowed, was introduced by BOSE, SHRIKHANDE & PARKER [4] and H. HANANI [9], and played important roles in their respective work on orthogonal Latin squares and BIBD's.

The purpose of this paper is to present some of the methods for constructing designs and to briefly indicate how PBD's have and can be used in the construction of related combinatorial structures. We also take this opportunity to add various remarks and indicate variations on proofs of known results. Many of the proofs, if not omitted, will be very brief.

A PBD$[K,v]$ is to be a PBD $(X,\mathcal{A})$ where $|X| = v$ and $|A| \in K$ for every $A \in \mathcal{A}$. Here K is a (finite or infinite) set of positive integers. For the case where K consists of a single positive integer k, we write B$[k,v]$ in place of PBD$[\{k\},v]$.

We observe that the existence of a PBD$[K,v]$ (with $v > 0$) implies

(i) $v \equiv 1 \pmod{\alpha(K)}$, and

(ii) $v(v-1) \equiv 0 \pmod{\beta(K)}$,

where $\alpha(K)$ is the greatest common divisor of the integers $\{k-1 : k \in K\}$ and

---

*) This research was supported in part by NSF Grant GP-28943 (O.S.U.R.F. Project No. 3228-A1).

$\beta(K)$ is the greatest common divisor of the integers $\{k(k-1) : k\in K\}$. Here (i) follows from the fact that the blocks containing a given point of a PBD partition the remaining v-1 points; and (ii) follows since the $\binom{v}{2}$ pairs of points are partitioned by the blocks.

The above conditions (i) and (ii) are "asymptotically sufficient" for the existence of a PBD[K,v]. The following theorem is proved in [24] using the methods discussed in this paper.

THEOREM 1.1. *Given* K, *there exists a constant* c_K *such that designs* PBD[K,v] *exist for all* $v \geq c_K$ *which satisfy the congruences* $v \equiv 1 \pmod{\alpha(K)}$ *and* $v(v-1) \equiv 0 \pmod{\beta(K)}$.

For example, PBD[{7,8,9},v] exist for all large integers v; PBD[{5,7},v] exist for all large odd integers v; and B[6,v] exist for all large $v \equiv 1,6,16$, or 21 (mod 30). Complete characterizations of the set $\mathbb{B}(K)$ of positive integers v for which there exist designs PBD[K,v] are known only for a few sets K (see section 5).

In section 2, we use difference methods and finite fields to construct BIBD's. As far as the author is aware, this provides the only known method of ascertaining for each k the existence of a design B[k,v] for some $v > k$.

In section 3 we attempt to communicate the flavor of some purely combinatorial techniques for constructing PBD's. We give some very general methods, but do not attempt to survey the tremendous variety of constructions which are known for, say, Steiner triple systems.

The application of pairwise balanced designs in the construction of related combinatorial systems is illustrated in several places. We mention in section 3 how PBD's were used in the construction of resolvable designs. In section 4, we point out the use of PBD's in the construction of certain quasigroups (Latin squares). The results on PBD's and the difference method are used in section 6 to obtain infinite classes of what we call edge-decompositions of complete graphs.

2. CONSTRUCTION OF DESIGNS BY DIFFERENCE METHODS

The "method of differences" introduced by R.C. BOSE [1] has been an effective method for the construction of designs B[k,v], especially for small values of k. We consider below the simplest case of the method.

Let G be an abelian group of order $v = k(k-1)t+1$. By a *simple difference family* D[k,v] in G, we mean a family $A_1, A_2, \ldots, A_t$ of k-subsets of G such that every non-zero group element occurs exactly once in the list of differences

$$(x-y : x,y \in A_i;\ x \neq y;\ i=1,2,\ldots,t).$$

This includes the case of planar difference sets ($t = 1$).

Examples include the D[3,13] ({1,3,9}, {2,6,5}) in Z_{13} and the D[4,25] ({(0,0), (1,0), (0,1), (2,2)}, {(0,0),(2,0), (0,2), (4,4)}) in $Z_5 \times Z_5$.

The existence of a simple difference family D[k,v] implies the existence of a design B[k,v]. For the design, take X = G and $\mathcal{A} = \{A_i + g : i=1,2,\ldots,t;\ g \in G\}$.

The existence of simple difference families D[3,v] in cyclic groups of order $v = 6t+1$ was established in 1939 by R. PELTESOHN [16], and a construction of R.C. BOSE [1], also in 1939, shows that D[3,v] always exist in elementary abelian groups of order $v = 6t+1$.

We conjecture that for each fixed k, simple difference families D[k,v] exist in all but finitely many abelian groups of orders $v = k(k-1)t+1$. The following theorem provides some evidence towards this conjecture.

THEOREM 2.1. *If* $q = k(k-1)t+1$ *is a prime power and*

$$q > e^{k^2+2k \log k},$$

then there exists a simple difference family D[k,q] *in the elementary abelian group of order* q, *and hence a design* B[k,q].

This theorem in a weaker form (with the bound $q > [k(k-1)/2]^{k(k-1)}$) was proved in [21]. While this improvement in the bound is significant, it surely is still very far from best possible. (Note that here e denotes the exponential $e = 2.718\ldots$, in distinction to its use in [21].)

The main idea of the proof is to exploit the multiplicative structure of finite fields to find simple difference families in their additive groups. Proposition 2.1 below reduces the problem to finding a single k-subset with a certain property. It remains to establish the existence of such k-subsets for large q, and we take this opportunity to give a proof which is more combinatorial in nature than that of [21] (i.e., we avoid the use of

character sums).

Let $q = mf+1$ be a prime power and let $F = GF(q)$ be the field with q elements. The cyclic multiplicative group of F has a unique subgroup C_0 of index m (and order $f = (q-1)/m$). The multiplicative cosets $C_0, C_1, \ldots, C_{m-1}$ of C_0 are the *cyclotomic classes* of index m. They evidently partition $F - \{0\}$.

PROPOSITION 2.1. *Let* $q = k(k-1)t+1$ *be a prime power and put* $m = \frac{1}{2}k(k-1)$. *If there exists a* k*-tuple* $(a_1, a_2, \ldots, a_k)$ *of elements of* $GF(q)$ *such that the* m *differences*

$$(a_j - a_i : 1 \le i < j \le k)$$

form a system of representatives for the cyclotomic classes $C_0, C_1, \ldots, C_{m-1}$ *of index* m *in* $GF(q)$, *then there exists a simple difference family* $D[k,q]$ *in the additive group of* $GF(q)$.

PROOF. Let $A = \{a_1, \ldots, a_k\}$. Since $2m$ divides $q-1$, -1 will belong to C_0. Let S be a system of representatives for the cosets of the factor group $C_0/\{1,-1\}$ (i.e., S consists of half of the elements of C_0, one element from each pair $\{x,-x\} \subseteq C_0$).

We claim that $(sA : s \in S)$ is a simple difference family, where $sA = \{sa_1, \ldots, sa_k\}$. To see this, we need only observe that for each i,j $(1 \le i < j \le k)$, we find among the differences from $(sA : s \in S)$

$$sa_j - sa_i \quad \text{and} \quad sa_i - sa_j, \quad s \in S,$$

or

$$(\pm s(a_j - a_i) : s \in S),$$

which exhaust precisely the cyclotomic class represented by $a_j - a_i$. □

Examples for $k = 4$ include $(0,1,3,24)$ in $GF(37)$ and $(0,1,5,11)$ in $GF(61)$.

For prime powers $q \equiv 1 \pmod{k(k-1)}$, let $N(k,q)$ denote the number of k-tuples $(a_1, \ldots, a_k)$ with the property required in proposition 2.1. We prove that $N(k,q) > 0$, i.e. that such k-tuples exist, for q sufficiently large with respect to k by showing that

$$N(k,q) = \frac{m!}{m^m} q(q-1)\ldots(q-k+1) + O(q^{k-\frac{1}{2}}) ,$$

where $m = \binom{k}{2}$. This is stated in a more exact form as theorem 2.2 below. The corollary 2.1 of this theorem, together with proposition 2.1, will complete the proof of theorem 2.1.

LEMMA 2.1. *Let* $\alpha_1,\alpha_2,\ldots,\alpha_n$ *be real numbers,*

$$\bar{\alpha} = \frac{1}{n}(\alpha_1+\ldots+\alpha_n)$$

their mean, and

$$V = \frac{1}{n}\sum_{i=1}^{n} (\alpha_i-\bar{\alpha})^2 = \frac{1}{n}(\sum_{i=1}^{n} \alpha_i^2) - \bar{\alpha}^2$$

their variance. If we put $N = \alpha_1+\ldots+\alpha_l$ *for some* l, $0 \le l \le n$, *then*

$$|N - l\bar{\alpha}|^2 \le l(n-l)V \le \frac{1}{4}n^2 V.$$

PROOF. It is sufficient to establish the inequality in the case $\bar{\alpha} = 0$. For this case, let $\underline{a}$ denote the vector $(\alpha_1,\ldots,\alpha_n)$, $\underline{u} = (1,1,\ldots,1)$, and $\underline{w} = (1,\ldots,1,0,\ldots,0)$ (ones in the first l coordinates, zeros in the last $n-l$). Then the dot product $\langle \underline{a},\underline{u}\rangle$ is 0 and by the Cauchy-Schwarz inequality,

$$|\langle \underline{a},\underline{w}\rangle| = |\langle \underline{a},\underline{w} - \beta\underline{u}\rangle| \le ||\underline{a}|| \cdot ||\underline{w} - \beta\underline{u}||$$

for any real β. Taking $\beta = (n-l)/n$, we obtain

$$(\alpha_1+\ldots+\alpha_l)^2 \le (\alpha_1^2+\ldots+\alpha_n^2)\,\frac{l(n-l)}{n} ,$$

which is the desired inequality. □

For the following lemma, we fix a prime power $q = mf+1$ and write $F = GF(q)$, $Z_m = \{0,1,\ldots,m-1\}$. C_i, $i \in Z_m$, are the cyclotomic classes of index m. For a set X, X^r denotes the set of all r-tuples $(x_1,\ldots,x_r)$ of elements of X, and $X^{(r)}$ denotes the subset of X^r consisting of $(x_1,\ldots,x_r)$ with $x_1,\ldots,x_r$ distinct. Thus if $|X| = n$, then $|X^r| = n^r$ and $|X^{(r)}| = n^{(r)} = n(n-1)(n-2)\ldots(n-r+1)$.

For $i_1,i_2,\ldots,i_r \in Z_m$ and distinct $a_1,a_2,\ldots,a_r \in F$, let

$E_{i_1 i_2 \ldots i_r}(a_1, a_2, \ldots, a_r)$

denote the number of field elements $x \in F$ such that $x-a_1 \in C_{i_1}$, $x-a_2 \in C_{i_2}, \ldots, x-a_r \in C_{i_r}$.

LEMMA 2.2. *The mean value of* $E_{(i)}(a) = E_{i_1 \ldots i_r}(a_1, \ldots, a_r)$ *(over the* $m^r q^{(r)}$ *choices of* $(i) = (i_1, \ldots, i_r) \in Z_m^r$ *and* $(a) = (a_1, \ldots, a_r) \in F^{(r)}$*) is* $(q-r)m^{-r}$, *and the variance* V_r *of these* $m^r q^{(r)}$ *quantities is*

$$V_r = \frac{q(q-1)}{m^r q^{(r)}} \left[\frac{q-m-1}{m}\right]^{(r)} + \frac{q-r}{m^r} - \left(\frac{q-r}{m^r}\right)^2 < \frac{q-r}{m^r} .$$

PROOF. It is immediate that for $(a) \in F^{(r)}$,

$$\sum_{(i)} E_{(i)}(a) = q-r ,$$

and then

$$\sum_{(i),(a)} E_{(i)}(a) = q^{(r+1)} .$$

So the mean is as claimed.

With the usual notation $E^{(2)} = E(E-1)$, $[E_{(i)}(a)]^{(2)}$ is the number of pairs $(x,y) \in F^{(2)}$ such that $x-a_j$ and $y-a_j$ both belong to the class C_{i_j} for $j=1,2,\ldots,r$. Then for fixed $(a) \in F^{(r)}$,

$$\sum_{(i)} [E_{(i)}(a)]^{(2)}$$

is the number of $(x,y) \in F^{(2)}$ such that $x-a_j$ and $y-a_j$ belong to the same cyclotomic class for $j=1,2,\ldots,r$; and

$$\sum_{(i),(a)} [E_{(i)}(a)]^{(2)}$$

counts the number of $(r+2)$-tuples $(a_1, \ldots, a_r; x, y)$ with $(a) \in F^{(r)}$, $(x,y) \in F^{(2)}$, and such that $x-a_j$ and $y-a_j$ are in the same class. For fixed $(x,y) \in F^{(2)}$, such an $(r+2)$-tuple is obtained by choosing $a_1, \ldots, a_r$ as distinct elements of the set $S(x,y)$ of c with $x-c$ and $y-c$ in the same class.

Now x-c and y-c are in the same class if and only if $x-c = b(y-c)$ for some $b \in C_0$. But for each of the $(q-m-1)/m$ elements $b \in C_0$, $b \neq 1$, there is a unique such solution c; that is, $|S(x,y)| = (q-m-1)/m$ (independent of x,y). So

$$\sum_{(i),(a)} [E_{(i)}(a)]^{(2)} = q(q-1)\,[\frac{q-m-1}{m}]^{(r)} .$$

The computation of the variance V_r is now straightforward, and the inequality $V_r < m^{-r}(q-r)$ is elementary. □

THEOREM 2.2. *Let* $m = \frac{1}{2}k(k-1)$. *Then for prime powers* $q = 2mt+1$,

$$|N(k,q) - \frac{m!}{m^m} q^{(k)}| < m^{\frac{1}{2}(k-1)} q^{k-\frac{1}{2}}.$$

PROOF. For $0 \le r \le k$, let M_r denote the set of $(a_1,\ldots,a_r) \in F^{(r)}$ such that the differences a_j-a_i, $1 \le i < j \le r$, represent $\binom{r}{2}$ distinct cyclotomic classes and write $M_r = |M_r|$. Thus $M_0 = 1$, $M_1 = q$, $M_2 = q(q-1)$, and $M_k = N(k,q)$. The members of M_{r+1} may be partitioned according to their first r coordinates (which determine a member of M_r) and we evidently have

$$M_{r+1} = \sum_{(a)\in M_r} E'(a) ,$$

where $E'(a_1,\ldots,a_r)$ is the number of x such that $(a_1,\ldots,a_r,x) \in M_{r+1}$. But clearly

$$E'(a_1,\ldots,a_r) = \sum E_{i_1\ldots i_r}(a_1,\ldots,a_r) ,$$

where the sum is extended over the $[m-\binom{r}{2}]^{(r)}$ choices of distinct $i_1,\ldots,i_r$ chosen from the set of $m-\binom{r}{2}$ elements $i \in Z_m$ for which the class C_i is *not* represented by a difference from $(a_1,\ldots,a_r)$.

In summary, M_{r+1} is a sum of $M_r[m-\binom{r}{2}]^{(r)}$ of the quantities $E_{(i)}(a)$. By lemmas 2.1 and 2.2,

$$|M_{r+1} - \frac{q-r}{m^r}[m-\tbinom{r}{2}]^{(r)} M_r| \le (\tfrac{1}{4}m^r q^{(r)} V_r)^{\frac{1}{2}} < (\tfrac{1}{4}m^r q^{2r+1})^{\frac{1}{2}}.$$

With

$$c_r = \frac{q-r}{m^r}\,[m-\tbinom{r}{2}]^{(r)}, \quad d_r = \tfrac{1}{2}m^{r/2}\,q^{r+\frac{1}{2}},$$

we have $|M_{r+1}-c_rM_r| < d_r$ for $r=0,1,\ldots,k-1$. Using the triangle inequality inductively, and $M_0 = 1$, we arrive at

$$|M_{r+1}-c_rc_{r-1}\cdots c_1c_0| < d_r+c_rd_{r-1}+c_rc_{r-1}d_{r-2}+\ldots+c_rc_{r-1}\cdots c_1d_0.$$

Now $c_r < q$ and

$$c_{k-1}c_{k-2}\cdots c_1c_0 = \frac{m!}{m^m}\,q^{(k)},$$

so

$$|M_k-\frac{m!}{m^m}\,q^{(k)}| < \tfrac{1}{2}q^{k-\frac{1}{2}}(m^{(k-1)/2}+m^{(k-2)/2}+\ldots+1) < m^{(k-1)/2}q^{k-\frac{1}{2}}. \quad \square$$

COROLLARY 2.1. *If* $q = k(k-1)t+1$ *is a prime power, then* $N(k,q) > 0$ *whenever*

$$q > e^{k^2}\,k^{2k}.$$

PROOF. By theorem 2.2,

$$N(k,q) > \frac{m!}{m^m}\,q^{(k)} - m^{\frac{1}{2}(k-1)}\,q^{k-\frac{1}{2}},$$

$$q^{\frac{1}{2}-k}\,N(k,q) > \frac{m!}{m^m}\,\sqrt{q}\,(\frac{q^{(k)}}{q^k}) - m^{\frac{1}{2}(k-1)},$$

where $m = k(k-1)/2$. Using the inequality $m!/m^m > e^{-m}$ (which is immediate on noticing that $m^m/m!$ is one of the terms in the power series expansion of e^m) and the (very poor) inequality $q^{(k)}/q^k > e^{-k/2}$ for q satisfying the hypothesis,

$$q^{\frac{1}{2}-k}\,N(k,q) > e^{-k^2/2}\,\sqrt{q} - k^k.$$

The assertion of the corollary is now clear. $\square$

3. CONSTRUCTION OF DESIGNS BY COMPOSITION METHODS

In this section we discuss a class of recursive methods for the construction of PBD's. We make no attempt to be complete, but just enumerate certain principles and illustrations that the author finds of interest.

A concept which has played an important role in the construction of BIBD's and sets of orthogonal Latin squares is that of a *group divisible design* (GDD). We use the term to mean a triple $(X,\mathcal{S},\mathcal{A})$ where (i) X is a set (of *points*), (ii) $\mathcal{S}$ is a class of non-empty subsets of X (called *groups*) which partition X, (iii) $\mathcal{A}$ is a class of subsets of X (called *blocks*), each containing at least two points, (iv) no block meets a group in more than one point, and (v) each pairset $\{x,y\}$ of points not contained in a group is contained in precisely one block.

At least in the case where all groups $G \in \mathcal{S}$ have size $|G| \geq 2$, a GDD can be thought of as a PBD $(X,\mathcal{S} \cup \mathcal{A})$ in which a class of blocks which partition X has been distinguished. But it seems important to make the distinction between PBD's and GDD's, as GDD's are clearly the right concept for the Fundamental Construction (F.C.) 3.1 below. We preface the F.C. with several remarks concerning the relation between PBD's and GDD's.

REMARK 3.1. If $\mathcal{S}$ consists of all singleton subsets of X, then $(X,\mathcal{A})$ is a PBD if and only if $(X,\mathcal{S},\mathcal{A})$ is a GDD.

REMARK 3.2. (ADJOINING AND DELETING POINTS). If $(X,\mathcal{S},\mathcal{A})$ is a GDD, we may *adjoin a point* $\theta \notin X$ to obtain a PBD $(X',\mathcal{A}')$ where

$$X' = X \cup \{\theta\} ,$$

$$\mathcal{A}' = \mathcal{A} \cup \{G \cup \{\theta\} : G \in \mathcal{S}\} .$$

Conversely, given a PBD $(X',\mathcal{A}')$ we may *delete a point* $\theta \in X'$ to obtain a GDD $(X,\mathcal{S},\mathcal{A})$, where

$$X = X' - \{\theta\} ,$$

$$\mathcal{S} = \{A - \{\theta\} : A \in \mathcal{A}', \theta \in A\} ,$$

$$\mathcal{A} = \{A : A \in \mathcal{A}', \theta \notin A\}.$$

By a *subdesign* of a PBD $(X,\mathcal{B})$, we mean a PBD $(Y,\mathcal{C})$ such that $Y \subseteq X$ and $\mathcal{C} \subseteq \mathcal{B}$. Evidently, the blocks $\mathcal{B}-\mathcal{C}$ cover exactly the pairs x,y of distinct

points of X with not both $x,y \in Y$. We admit $|Y| \leq 1$, in which case $\mathcal{C}$ is empty.

REMARK 3.3. (ADJOINING SUBDESIGNS). Let $(X,S,\mathcal{A})$ be a GDD and F a set, $F \cap X = \emptyset$. Let $(F,\mathcal{D})$ be a PBD, and for each group $G \in S$, let $(G \cup F, \mathcal{B}_G)$ be a PBD containing a PBD $(F,\mathcal{C}_G)$ as a subdesign. Then with

$$X' = X \cup F ,$$

$$\mathcal{A}' = \mathcal{A} \cup \mathcal{D} \cup \left(\bigcup_{G \in S} \mathcal{B}_G - \mathcal{C}_G\right) ,$$

$(X',\mathcal{A}')$ is a PBD.

REMARK 3.4. (BREAKING UP BLOCKS). Let $(X,\mathcal{A})$ be a PBD and for each block $A \in \mathcal{A}$, let $(A,\mathcal{B}_A)$ be a PBD. Then with

$$\mathcal{B} = \bigcup_{A \in \mathcal{A}} \mathcal{B}_A ,$$

$(X,\mathcal{B})$ is a PBD.

THE FUNDAMENTAL CONSTRUCTION (F.C.) 3.1. Let $(X,S,\mathcal{A})$ be a "master" GDD and let a positive integral weight s_x be assigned to each point $x \in X$. Let $(S_x : x \in X)$ be pairwise disjoint sets with $|S_x| = s_x$. With the notation $S_Y = \bigcup_{x \in Y} S_x$ for $Y \subseteq X$, put

$$X^* = S_X$$

$$S^* = \{S_G : G \in S\} .$$

For $A \in \mathcal{A}$, we have a natural partition $\pi_A = (S_A,\{S_x : x \in A\})$; we suppose that for each block $A \in \mathcal{A}$, a GDD

$$(S_A ,\{S_x : x \in A\}, \mathcal{B}_A)$$

is given, and put

$$\mathcal{A}^* = \bigcup_{A \in \mathcal{A}} \mathcal{B}_A .$$

Then $(X^*,S^*,\mathcal{A}^*)$ is a GDD.

We point out that in view of remark 3.1, remark 3.4 is a special case

of the F.C. 3.1 (where the master GDD has groups of size 1 and we weight each of its points with 1).

We give below some of the neater corollaries of our remarks and the F.C. In the various applications here and in the following section, all points of the master GDD in the F.C. will be weighted equally. But instances of non-uniform weighting were essential for the proof of theorem 1.1.

A GDD $(X,\mathcal{S},\mathcal{A})$ will be said to have block sizes from K and *type* $(1^{l_1}2^{l_2}3^{l_3}\ldots)$ when $|A| \in K$ for each $A \in \mathcal{A}$ and $(1^{l_1}2^{l_2}3^{l_3}\ldots)$ is the partition of the integer $|X|$ arising from the partition $(X,\mathcal{S})$ of the set X, i.e. there are l_i groups of size i, i = 1,2,3,... . We say a class $\mathcal{B}$ of sets is k-*uniform* when $|B| = k$ for each $B \in \mathcal{B}$.

Transversal designs TD(k,n) are GDD's with type (n^k), i.e. k groups of size n, and block size k. The existence of a TD(k,n) is equivalent to the existence of a set of k-2 mutually orthogonal Latin squares [8].

THEOREM 3.1. (MACNEISH). *If there exists a* TD(k,n) *and a* TD(k,m), *then there exists a* TD(k,mn).

PROOF. Take the TD(k,n) as the master GDD in the F.C., and weight each of its points with m. The type of each partition π_A is then (m^k), and as there exists a TD(k,m), we may choose $\mathcal{B}_A$ to be k-uniform. The F.C. then produces a-TD(k,mn). □

From the existence of TD(q+1,q) for prime powers q (obtainable by deleting a point from a $B[q+1, q^2+q+1]$ by remark 3.2) follows

THEOREM 3.2. *If* $n = p_1^{\alpha_1}p_2^{\alpha_2}\ldots p_r^{\alpha_r}$ *is the factorization of* $n > 1$ *into powers of distinct primes, then there exists a* TD(k,n) *whenever* $k \le 1 + \min_{1\le i\le r} p_i^{\alpha_i}$.

The following theorem was first proved in the case k = 3 by E.H. MOORE [15] and in 1893. The generalization was pointed out to the author by D.K. RAY-CHAUDHURI.

THEOREM 3.3. *If*

(i) *there exists a* B[k,u] *containing a* B[k,w] *as a subdesign* (w = 0,1, *or* k *is permitted), and*

(ii) *there exists a* TD(k,u-w),

then the existence of a B[k,v] *implies the existence of a* B[k,v(u-w)+w].

PROOF. As the master GDD in the F.C., take a B[k,v] with all singletons as groups, and weight each point with u-w. The type of each partition π_A is

$((u-w)^k)$ and $\mathcal{B}_A$ may be taken to be k-uniform by (ii). The F.C. produces a GDD of type $((u-w)^v)$, to which we may adjoin a subdesign on w points by (i) and remark 3.3 to obtain a B[k,v(u-w)+w]. □

THEOREM 3.4. *If*

(i) *there exists a* B[k,u] *containing a* B[k,w] *as a subdesign* $w \geq 1$, *and*

(ii) *there exists a* TD(k,(u-w)/(k-1)),

then the existence of a B[k,v] *implies the existence of a* B[k,(v-1)(u-w)/(k-1)+w].

PROOF. Delete a point (remark 3.2) from a B[k,v] to obtain a GDD with group type $((k-1)^r)$, $r = (v-1)/(k-1)$, and block size k. Weight each point with (u-w)/(k-1) (which is an integer since $u \equiv w \equiv 1 \pmod{k-1}$). By (ii), $\mathcal{B}_A$ may be chosen k-uniform. The F.C. produces a GDD of type $((u-w)^r)$ to which we adjoin a subdesign on w points by (i) to obtain a B[k,r(u-w)+w]. □

The next theorem is an elegant and powerful construction due to HANANI [9]. For each $k \geq 2$, we let R_k denote the set of positive integers r for which there exists a B[k,r(k-1)+1].

THEOREM 3.5. *If there exists a* PBD[R_k,v], *then* $v \in R_k$.

PROOF. By remark 3.2, R_k is exactly the set of positive integers r for which there exists a GDD of type $((k-1)^r)$ with block size k.

As the master GDD in the F.C., take a PBD[R_k,v] considered as a GDD of type (k^v) with block sizes from R_k. Weight each point with k-1. The type of a partition π_A is $((k-1)^r)$ where $|A| = r \in R_k$, and so $\mathcal{B}_A$ may be chosen k-uniform. The F.C. produces a GDD of type $((k-1)^v)$, block size k, which by the observation of the previous paragraph means $v \in R_k$. □

A PBD (X,A) is said to be *resolvable* iff there exists a partition

$$A = A_1 \cup A_2 \cup \ldots \cup A_r$$

such that each set A_i of blocks is a partition of X (a *parallel class* of blocks). If we let R_k^* denote the set of positive integers r for which there exists a $B^*[k,r(k-1)+1]$ (i.e., a resolvable B[k,r(k-1)+1]), then we have

THEOREM 3.6. *If there exists a* PBD[R_k^*,v], *then* $v \in R_k^*$.

This theorem, due to RAY-CHAUDHURI and the author, may be found in [17]. We remark only that it can be proved with the same construction as in theorem 3.5, and that it was an important step in the proofs that $B^*[3,6t+3]$ and $B^*[4,12t+4]$ exist for all t (see [17,13]) and that $B^*[k,k(k-1)t+k]$ exist for all t sufficiently large with respect to k (see [18]).

Resolvable designs $B^*[k,v]$ (with fixed parallelism, i.e. partition of the blocks into parallel classes) are also known as Sperner spaces.

Instances and/or extensions of the constructions discussed in this section can be found in [3,9,11,12,17,22,23].

We conclude this section with two remarks to be used in section 5. We do not attempt to state these in the most general form.

REMARK 3.5. (COMPLETION). If $(X,\mathcal{A})$ is a resolvable $B[k,v]$, say $\mathcal{A} = \mathcal{A}_1 \cup \ldots \cup \mathcal{A}_r$ is a partition of $\mathcal{A}$ into parallel classes ($r = (v-1)/(k-1)$), then we may *complete* $(X,\mathcal{A})$ to a $PBD[\{k+1,r\},v+r]$ $(X',\mathcal{A}')$ by adding r "points at infinity" $\theta_1,\ldots,\theta_r$ and a "block at infinity" $B = \{\theta_1,\ldots,\theta_r\}$. That is, we put

$$X' = X \cup \{\theta_1,\ldots,\theta_r\} ,$$

$$\mathcal{A}' = \{B\} \cup \{A \cup \{\theta_i\} : A \in \mathcal{A}_i;\ i=1,2,\ldots,r\} .$$

For $1 \le r$, we may partially complete $(X,\mathcal{A})$ by adding a block of size 1 "at infinity" (adjoining "points at infinity" only to 1 of the parallel classes $\mathcal{A}_i$) to obtain a $PBD[\{k,k+1,1\},v+1]$.

REMARK 3.6. (TRUNCATION). Given a transversal design $TD[k+1,t]$, we obtain a GDD of type (s^1t^k) with block sizes from $\{k,k+1\}$ by deleting t-s points from one of the groups (and from the blocks which contain them) of the transversal design.

4. A CLOSURE OPERATION

Given a set K (finite or infinite) of positive integers, we denote by $\mathbb{B}(K)$ the set of positive integers v for which there exists a $PBD[K,v]$. The mapping $K \to \mathbb{B}(K)$ is a *closure operation* on the subsets of the positive integers; that is, it enjoys the properties

(i) $K \subseteq \mathbb{B}(K)$,

(ii) $K_1 \subseteq K_2 \Rightarrow \mathbb{B}(K_1) \subseteq \mathbb{B}(K_2)$,

(iii) $\mathbb{B}(\mathbb{B}(K)) = \mathbb{B}(K)$.

These are easily verified. (Property (iii) is a consequence of remark 3.4.)

We call a set K of positive integers *PBD-closed* (or simply *closed*) when $\mathbb{B}(K) = K$. Theorem 1.1 asserts that every closed set K contains all sufficiently large integers v with $v \equiv 1 \pmod{\alpha(K)}$ and $v(v-1) \equiv 0 \pmod{\beta(K)}$. Thus there are only countably many closed sets.

Using the fact that the greatest common divisor of any set is equal to the g.c.d. of a finite subset, a consequence of theorem 1.1 is (see [23])

THEOREM 4.1. *If* K *is a closed set, then there exists a finite subset* $J \subseteq K$ *such that* $K = \mathbb{B}(J)$.

Examples of closed sets seem to arise naturally in certain existence problems. We give several such examples below. The proof that a set defined combinatorially is closed invariably involves a construction (as in remark 3.4, which shows that $\mathbb{B}(K)$ is closed for any set K). Of course, the construction contains far more information than simply the assertion that a set is closed.

Theorems 3.5 and 3.6 can be rephrased as

EXAMPLE 4.1. For any $k \geq 2$, the sets R_k and R_k^* are closed.

To apply theorem 1.1 to a closed set K, it is necessary to know something about the parameters $\alpha(K)$ and $\beta(K)$. This involves exhibiting some elements of K. For if we know $k \in K$, then already we can say $\beta(K)$ divides $k(k-1)$ and hence all large $v \equiv 1 \pmod{k(k-1)}$ belong to K. More generally, if we know $\beta(K)$ divides b and $u \in K$, K a closed set, then K contains all large $v \equiv u \pmod{b}$.

We point out some simple arithmetic examples of closed sets.

EXAMPLE 4.2. Let a be a positive integer. Then $H^a = \{v : v \equiv 1 \pmod{a}\}$ is closed.

For a clearly divides $\alpha(H_a)$; hence $v \in \mathbb{B}(H^a)$ implies $v \in H^a$.

EXAMPLE 4.3. Let b be a positive integer. Then $H_b = \{v : v(v-1) \equiv 0 \pmod{b}\}$ is closed.

For b divides $\beta(H_b)$; hence $v \in \mathbb{B}(H_b)$ implies $v \in H_b$.

EXAMPLE 4.4. Let m be a positive integer. Then $\{1\} \cup \{v : v \geq m\}$ is closed.

This is easily seen. Since the intersection of closed sets is closed, we have

EXAMPLE 4.5. Let a,b,m be positive integers. Then

$$H_b^a = \{v : v \equiv 1 \pmod a, v(v-1) \equiv 0 \pmod b\}$$

and

$$H_b^a \cap \{v : v = 1 \text{ or } v \geq m\} \text{ are closed.}$$

We remark that $\alpha(H_b^a) = a$ and $\beta(H_b^a) = b$ if and only if a divides b, b is even, and b/a is relatively prime to a (allowing a = b = 0) [23].

The next several examples of closed sets arise from the following observation that PBD's can be used to combine idempotent quasigroups to form another, larger, idempotent quasigroup.

A *quasigroup* on a finite set X is a binary operation Q: X×X → X which satisfies both cancellation laws, i.e. the values of any two of x,y,z ∈ X uniquely determine the value of the third so that the equation xQy = z is valid. The quasigroup Q is *idempotent* when xQx = x for all x ∈ X.

Let (X,A) be a PBD and for each block A ∈ A, let Q_A be an idempotent quasigroup on A. Then define Q on X by xQx = x, and for distinct x,y ∈ X, $xQy = xQ_Ay$ where A is the unique block of A containing x and y. It is easily checked that Q is an idempotent quasigroup on X.

EXAMPLE 4.6. Given k, let L_k' denote the set of positive integers n for which there exists a set of k mutually orthogonal Latin squares of order n which admit a common transversal (we may suppose, e.g., that all symbols occur on the diagonal of each square). Then L_k' is a closed set.

This is an observation of BOSE, SHRIKHANDE & PARKER [4] and was instrumental in their disproof of EULER's conjecture (see also [8,14]). We find it convenient here to explain their construction in terms of quasigroups.

Two quasigroups Q_1, Q_2 on the same set X are *orthogonal* iff for any a,b ∈ X, the system of equations

$$xQ_1y = a, \quad xQ_2y = b$$

has a unique simultaneous solution (x,y). The assertion n ∈ L_k' is equivalent to the existence of k mutually orthogonal idempotent quasigroups of order n.

Now given n ∈ $\mathbb{B}(L_k')$, there exists a PBD (X,A) with |X| = n and |A| ∈ L_k' for all A ∈ A. For each block A, we can find k mutually orthogonal idempotent quasigroups $Q_A^{(1)}, \ldots, Q_A^{(k)}$ on A. Then the above construction

produces idempotent quasigroups $Q^{(1)},\ldots,Q^{(k)}$ on X, which are readily seen to be mutually orthogonal. Hence $n \in L_k'$ and we have shown that L_k' is closed.

There is no point in applying theorem 1.1 to the existence problem for orthogonal Latin squares, since these results are used heavily in the proof of theorem 1.1. However, one can easily see that the number of non-isomorphic sets of k mutually orthogonal Latin squares of order n goes to infinity with n (use theorem 4.1 in a manner analogous to the proof of theorem 2 in [24]).

A quasigroup Q is *self-orthogonal* when it is orthogonal to its transpose Q' defined by $xQ'y = yQx$.

EXAMPLE 4.7. The set S of positive integers n for which there exists a self-orthogonal quasigroup of order n is closed.

The same construction works, as noticed by J.F. LAWLESS [14]. Using $S = \mathbb{B}(S)$ and some special constructions, BRAYTON, COPPERSMITH & HOFFMAN [5] have recently shown that S contains all positive integers except 2,3 and 6. Again, we claim that the number of non-isomorphic such quasigroups tends to infinity with n.

A *Room pair* of quasigroups is a pair Q_1, Q_2 of commutative idempotent quasigroups on the same set X such that for any $a,b \in X$, there is at most one unordered pair $\{x,y\}$ with $xQ_1y = a$ and $xQ_2y = b$.

EXAMPLE 4.8. The set R of positive integers n such that there exists a Room pair of quasigroups of order n is closed.

This was first noticed by J.F. LAWLESS [14]. The construction is also discussed in [19].

We conclude our remarks on quasigroups by observing that GDD's can be used to construct (not necessarily idempotent) quasigroups: if each block A of a GDD (X,S,A) is equipped with an idempotent quasigroup Q_A and each group G is equipped with any quasigroup Q_G, then there is a natural quasigroup Q on X, where $xQx = xQ_Gx$ for $x \in G$. When S consists of all singleton subsets of X, this construction degenerates into the basic construction with PBD's (cf. remark 3.1).

We introduce designs $PBD_\lambda[K,v]$, where each pair of distinct points is contained in exactly λ blocks, just for the purpose of stating

EXAMPLE 4.9. For any set K of positive integers and a positive integer λ, the set $\mathbb{B}_\lambda(K) = \{v : \text{there exists a } PBD_\lambda[K,v]\}$ is a closed set.

Other examples of closed sets are

EXAMPLE 4.10. The set of positive integers r for which there exists a reverse Steiner triple system of order 2r+1 (J. DOYEN [6]).

EXAMPLE 4.11. The set of v for which there exists a pair of orthogonal Steiner triple systems of order v (J.F. LAWLESS [14]).

EXAMPLE 4.12. The set of integers n for which there exists a GDD of type (m^n) and block sizes from K (see [22,20]).

One further example is given in section 6.

5. GENERATING CLOSED SETS

We observe that each closed set K has a unique minimal generating set. Let us call an element $x \in K$ *essential* in K iff $x \notin \mathbb{B}(K - \{x\})$, or equivalently, $x \notin \mathbb{B}(\{y \in K : y < x\})$. Letting E_K denote the set of all essential elements of K, we have

PROPOSITION 5.1. *Let J be a subset of a closed set K. Then* $\mathbb{B}(J) = K$ *if and only if* $E_K \subseteq J$.

PROOF. Clearly, $\mathbb{B}(J) = K$ implies $E_K \subseteq J$. We claim now that $\mathbb{B}(E_K) = K$. If not, let x be the least element of $K - \mathbb{B}(E_K)$. Then x is not essential, so $x \in \mathbb{B}(\{y \in K : y < x\}) \subseteq \mathbb{B}(\mathbb{B}(E_K)) = \mathbb{B}(E_K)$, a contradiction. □

Note that theorem 4.1 implies E_K is finite.

We may ask for the set of essential elements, or at least a reasonably small generating set, for the arithmetic examples H_b^a of section 4. HANANI [9,10] establishes that

$$H_3^1 = \{0,1 \pmod 3\} = \mathbb{B}\{3,4,6\} ,$$

$$H_4^1 = \{0,1 \pmod 4\} = \mathbb{B}\{4,5,8,9,12\} ,$$

$$H_5^1 = \{0,1 \pmod 5\} = \mathbb{B}\{5,6,10,11,15,16,20,35,36,40,70,71,75,76\} ,$$

as preliminary results before proving

$H_6^2 = \{1,3 \pmod 6\} = IB\{3\}$,

$H_{12}^3 = \{1,4 \pmod{12}\} = IB\{4\}$,

$H_{20}^4 = \{1,5 \pmod{20}\} = IB\{5\}$.

The reader may check that the essential elements of H_3^1 and H_4^1 are exactly those listed. We can improve the result for H_5^1 to

PROPOSITION 5.2. $H_5^1 = IB\{5,6,10,11,16,20,35,40\}$.

PROOF. We assume HANANI's result above and proceed to show that 36,70,71, 75,76 are not essential in H_5^1.

R.C. BOSE [2] has shown the existence of resolvable $B[q+1, q^3+1]$ for all prime powers q. In particular, we have the existence of a resolvable B[5,65], which we may partially complete (remark 3.5) by adding a block at infinity of size 5,6,10,11, respectively, to find $70,71,75,76 \in IB\{5,6,10,11\}$.

To see 36 is non-essential in H_5^1, we remove two intersecting lines (and their points) from a B[7,49] (affine plane of order 7) to obtain a PBD on 36 points with 18 blocks of size 6 and 36 of size 5 (i.e., $36 \in IB\{5,6\}$). □

The author does not know whether 35 and 40 are essential in H_5^1.

HANANI's paper [9] also exhibits finite generating sets for the closed sets $\{1\} \cup \{v : v \geq k\}$ for k = 3,4,5. And the result $\{1\} \cup \{v : v \geq k\} = IB\{k,k+1,k+2,\ldots,kq_k-1\}$ where q_k is the smallest prime power such that $q_k \geq 2k-1$, can be found in [12].

THEOREM 5.1.

(i) $H_2^2 = \{1 \pmod 2\} = IB\{3,5\}$,

(ii) $H_3^3 = \{1 \pmod 3\} = IB\{4,7,10,19\}$,

(iii) $H_4^4 = \{1 \pmod 4\} = IB\{5,9,13,17,29,33,49,57,89,93,129,137\}$.

LEMMA 5.1. *If* k *and* k+1 *are prime powers, the existence of a* GDD *on* v *points with block sizes from* {k+1,k+2} *and at least two groups implies that* vk+1 *is not essential in* $H_k^k = \{1 \pmod k\}$.

PROOF. We take such a GDD as the master GDD in the Fundamental Construction 3.1. Each point is to be weighted with k. The type of any partition π_A is (k^{k+1}) or (k^{k+2}). The classes $\mathcal{B}_A$ may be taken to be (k+1)-uniform as GDD's

with these types and block size k+1 may be obtained by deleting points (remark 3.2) from $B[k+1, k^2+k+1]$ and $B[k+1,(k+1)^2]$, respectively. The F.C. produces a GDD on vk points with groups of size divisible by k, and blocks of size k+1. Adjoining a point by remark 3.2, we obtain a PBD on vk+1 points with block sizes from $H_k^k = \{1 \pmod k\}$ (and less than vk+1). This proves the lemma. □

PROOF OF THEOREM 5.1. If there exists a TD(k+2,t), remark 3.6 shows that GDD's with v points satisfying the hypothesis of lemma 5.1 exist for $(k+1)t \le v \le (k+2)t$. With this observation and theorem 3.2, it can be seen (we omit the details) that with E_k denoting the essential elements of H_k^k,

$$E_2 \subseteq \{3,5,11,13,15,17\} ,$$

$$E_3 \subseteq \{4,7,10\} \cup \{19,22,\ldots,43,46\} \cup \{79,82\} ,$$

$$E_4 \subseteq \{5,9,13,17\} \cup \{29,33,\ldots,93,97\} \cup \{125,129,133,137\} .$$

We proceed to show that some of the indicated integers are not essential in the respective sets H_k^k.

Case k = 2. 13 and 15 belong to $IB\{3\}$ and hence are certainly non-essential in H_2^2. Completing a resolvable design $B^*[2,6]$ by remark 3.5 shows $11 \in IB\{3,5\}$. Deleting a point (remark 3.2) from a B[3,9] gives a GDD satisfying the hypothesis of lemma 5.1; hence 17 = $= 2\cdot 8+1$ is non-essential in H_2^2. Thus $E_2 \subseteq \{3,5\}$.

Case k = 3. $25,28,37,40 \in IB\{4\}$. Completing $B^*[3,15]$ and $B^*[3,21]$ shows $22 \in IB\{4,7\}$ and $31 \in IB\{4,10\}$. We show 43,46,79,82 are non-essential in H_3^3 by exhibiting GDD's on 14,15,26,27 points, respectively, with blocks of size 4 and using lemma 5.1. For 15 and 27, delete points (remark 3.2) from B[4,16] and B[4,28]. For 14, take the 14 points to be Z_{14}, groups $\{i,i+7\}$ $(i=0,1,\ldots,6)$, and blocks $\{2+i,4+i,8+i,7+i\}$ $(i=0,1,\ldots,13)$. For 26, take points Z_{26}, groups $\{i,i+13\}$ $(i=0,1,\ldots,12)$, and blocks $\{2+i,6+i,18+i, 13+i\}$, $\{4+i,12+i,10+i,13+i\}$ $(i=0,1,\ldots,25)$. Finally, $34 \in IB\{4,7,10\}$ follows from lemma 5.2 below (q = 3). We have proved $E_3 \subseteq \{4,7,10,19\}$ (and it is easily checked that equality holds).

Case k = 4. $41,45,61,65,81,125 \in IB\{5\}$ (cf.[9]) and $73 \in IB\{9\}$ (projective plane of order 8), hence these values are non-essential in H_4^4. Completing (by remark 3.5) resolvable designs $B^*[4,v]$ for v = $= 28,40,52,100$ (which exist [13]), we see that 37,53,69,133 are

non-essential in H_4^4. Deleting a point from a B[5,25] gives a GDD on 24 points and lemma 5.1 shows $97 = 4\cdot24+1$ is non-essential. Finally, lemma 5.2 below ($q = 4$) shows $77 \in \mathbb{B}\{5,13,17\}$. We have proved that all $x \equiv 1 \pmod 4$ are non-essential in H_4^4 with the possible exceptions of $x = 5,9,13,17,29,33,49,57,89,93,129,137$. □

LEMMA 5.2. *If* q *is a prime power, then* $q^3+q^2-q+1 \in \mathbb{B}\{q+1,q^2-q+1,q^2+1\}$.

PROOF. Let Π be a projective plane of order q^2 with a Baer subplane Π_0 (of order q). Let θ be a point of Π_0, $L_0,L_1,\ldots,L_q$ the lines of Π which contain θ and belong to the subplane Π_0, and L^* another line of Π containing θ. Let L_i' denote the set of the q^2-q points of L_i which do not belong to Π_0. The set $X = L^* \cup (\cup_{i=0}^{q} L_i')$ of $(q^2+1) + (q+1)(q^2-q)$ points of Π, together with the non-trivial intersections of lines of Π with X, provides a PBD with one block (namely L^*) of size q^2+1, $q+1$ blocks (namely, $L_i' \cup \{\theta\}$, $i=0,1,\ldots q$) of size q^2-q+1, and the remaining blocks have size $q+1$. This last assertion follows from the fact that each line of Π contains exactly one or $q+1$ points of Π_0. □

We indicate how theorem 5.1 (ii) can be used to derive two known results.

Kirkman Designs $B^*[3,v]$: Simple direct constructions [17] show that $B^*[3,v]$ exist for $v = 9,15,21,39$, i.e. $\{4,7,10,19\} \subseteq R_3^*$. By theorems 5.1 (ii) and 3.6, $H_3^3 = \mathbb{B}\{4,7,10,19\} \subseteq \mathbb{B}(R_3^*) = R_3^*$; which means that resolvable designs $B^*[3,6t+3]$ exist for all t.

Designs $B_2[4,3t+1]$ ($\lambda = 2$): If we establish that designs $B_2[4,v]$ exist for $v = 4,7,10,19$, it follows from theorem 5.1 (ii) and example 4.9 that $B_2[4,3t+1]$ exist for all t.

6. EDGE-DECOMPOSITIONS OF COMPLETE GRAPHS

Let Γ be a graph. By an *edge-decomposition* (or simply a *decomposition*) of a graph Γ^* into Γ-graphs we mean a set $\Gamma_1,\Gamma_2,\ldots,\Gamma_t$ of subgraphs of Γ^*, each isomorphic to Γ, such that each edge of Γ^* occurs in exactly one of the subgraphs Γ_i.

A design $B[k,v]$ may be considered as a decomposition of the complete graph on v vertices into k-cliques (complete graphs on k vertices). (More generally, a PBD can be viewed as a decomposition of a complete graph into cliques; and a GDD should perhaps be viewed as a decomposition of a complete

multipartite graph into cliques.)

Given a simple graph Γ, we denote by K_Γ the set of positive integers v for which the complete graph on v vertices admits a decomposition into Γ-graphs. The purpose of this last section is to provide a proof for

THEOREM 6.1. *Let Γ be a simple graph with* m *edges. Then K_Γ is a closed set and $\beta(K_\Gamma) = 2m$.*

That K_Γ is closed is easily verified; a consequence of theorem 6.1 and theorem 1.1 is that K_Γ contains all sufficiently large integers $v \equiv 1 \pmod{2m}$.

We observe that if $v \in K_\Gamma$, then surely the number m of edges of Γ divides $\binom{v}{2}$, i.e. $v(v-1) \equiv 0 \pmod{2m}$. Thus 2m is a divisor of $\beta(K_\Gamma)$. We show below that K_Γ contains the set Q of sufficiently large prime powers $q \equiv 1 \pmod{2m}$ (lemma 6.1), and observe that $\beta(Q) = 2m$ (lemma 6.2). $Q \subseteq K_\Gamma$ implies $\beta(K_\Gamma)$ divides $\beta(Q) = 2m$, and the proof of theorem 6.1 will then be complete. □

LEMMA 6.1. *If Γ is a simple graph with* m *edges and* k *vertices, then the complete graph on* q *vertices can be decomposed into Γ-graphs whenever* q *is a prime power of the form* $q = 2mt+1$ *and* $q > m^{k^2}$.

PROOF. Let $C_0, C_1, \ldots, C_{m-1}$ be the cyclotomic classes of index m in GF(q) and let S_i be the set of pairs {x,y} of distinct field elements such that $x-y \in C_i$ (since $-1 \in C_0$ in our case, x-y and y-x belong to the same class).

Theorem 3 of [21] asserts that if $q > m^{k^2}$, then for any choice of $l_{ij} \in \{0,1,\ldots,m-1\}$ there exist field elements $a_1, a_2, \ldots, a_k$ such that $\{a_i, a_j\} \in S_{l_{ij}}$ for all i,j ($1 \le i < j \le k$). (This can also be proved with lemmas 2.1 and 2.2). If we think of the edges of S_i as being colored with color i, then the claim is that all possible m-colorings are found among the edge colorings induced on k element subsets. So surely we may find a subgraph Γ_0 of the complete graph with vertex set GF(q) which is isomorphic to Γ and such that its m edges receive distinct colors.

Let S be a system of representatives for the cosets of the factor group $C_0/\{1,-1\}$. Then the set (not necessarily group) of permutations $\{x \to ax+b : a \in S, b \in GF(q)\}$ of GF(q) is sharply transitive on the edges of color i for each i. Applying these permutations to Γ_0 produces a set of isomorphic subgraphs which partition the edges of the complete graph on GF(q). □

LEMMA 6.2. *Let* m *and* c *be positive integers and let* Q *be the set of prime powers* q *such that* $q \equiv 1 \pmod{2m}$ *and* $q > c$. *Then* $\beta(Q) = 2m$.

PROOF. Clearly $\beta(Q)$ is divisible by 2m, say $\beta(Q) = 2mt$. We claim $t = 1$. If not, let p be a prime divisor of t. Now for at least one choice of sign, 2mp and $2m(p\pm1) + 1$ are relatively prime, and by DIRICHLET's theorem on primes in arithmetic progressions, there exists a prime $q = 2mpl + 2m(p\pm1)+1 > \max(c,\beta(Q))$. Then $q \in Q$, so $\beta(Q)$ divides $q(q-1)$. Since q is prime and larger than $\beta(Q)$, $\beta(Q) = 2mt$ divides $q-1$; hence 2mp divides $q-1 = 2m(pl+p\pm1)$. This contradiction shows that t has no prime divisors and completes the proof of the lemma. □

We close by remarking that theorem 6.1 can be used to prove a recent theorem of B. GANTER [7]. A partial PBD[K,v] is a system $(X,\mathcal{A})$ such that $|X| = v$, $|A| \in K$ for all $A \in \mathcal{A}$, and each pair of distinct points $x,y \in X$ is contained in at most one block $A \in \mathcal{A}$. GANTER's theorem asserts that for any finite partial PBD[K,v], there exists a finite PBD[K,v^*] $(X^*,\mathcal{A}^*)$ such that $X \subseteq X^*$ and $\mathcal{A} \subseteq \mathcal{A}^*$.

Now given a partial PBD $(X,\mathcal{A})$, we may consider the graph Γ with vertex set X and edge set consisting of those pairs {x,y} of points which do occur in some block $A \in \mathcal{A}$. If the complete graph with vertex set X^* can be decomposed into Γ-graphs, then it is clear that we may find a PBD $(X^*,\mathcal{A}^*)$ where $\mathcal{A}^*$ is a union of isomorphic copies of $\mathcal{A}$.

REFERENCES

[1] BOSE, R.C., *On the construction of balanced incomplete block designs*, Annals of Eugenics, 9 (1939) 353-399.

[2] BOSE, R.C., *On the application of finite projective geometry for deriving a certain series of balanced Kirkman arrangements*, Calcutta Math. Soc. Golden Jubilee Vol., 1959, pp. 341-354.

[3] BOSE, R.C. & S.S. SHRIKHANDE, *On the composition of balanced incomplete block designs*, Canad. J. Math., 12 (1960) 177-188.

[4] BOSE, R.C., S.S. SHRIKHANDE & E.T. PARKER, *Further results on the construction of mutually orthogonal Latin squares and the falsity of a conjecture of Euler*, Canad. J. Math., 12 (1960) 189-203.

[5] BRAYTON, R.K., D. COPPERSMITH & A.J. HOFFMAN, *Self-orthogonal Latin squares of all orders $n \neq 2,3,6$*, Bull. Amer. Math. Soc., 80 (1974) 116-118.

[6] DOYEN, J., *A note on reverse Steiner triple systems*, Discrete Math., 1 (1971-72) 315-319.

[7] GANTER, B., *Partial pairwise balanced designs*, Technische Hochschule Darmstadt Preprint No. 99, November, 1973.

[8] HALL JR., M., *Combinatorial theory*, Blaisdell, Waltham, Mass., 1967.

[9] HANANI, H., *The existence and construction of balanced incomplete block designs*, Ann. Math. Statist., 32 (1961) 361-386.

[10] HANANI, H., *A balanced incomplete block design*, Ann. Math. Statist., 36 (1965) 711.

[11] HANANI, H., *On balanced incomplete block designs with blocks having five elements*, J. Combinatorial Theory A, 12 (1972) 184-201.

[12] HANANI, H., *On balanced incomplete block designs and related designs*, to appear.

[13] HANANI, H., D.K. RAY-CHAUDHURI & R.M. WILSON, *On resolvable designs*, Discrete Math., 3 (1972) 343-357.

[14] LAWLESS, J.F., *Pairwise balanced designs and the construction of certain combinatorial systems*, in: Proceedings of the Second Louisiana Conference on Graph theory, Combinatorics and Computing, 1971.

[15] MOORE, E.H., *Concerning triple systems*, Math. Ann., 43 (1893) 271-285.

[16] PELTESOHN, R., *Eine Lösung der beiden Heffterschen Differenzenprobleme*, Compositio Math., 6 (1939) 251-257.

[17] RAY-CHAUDHURI, D.K. & R.M. WILSON, *Solution of Kirkman's school girl problem*, in: Proceedings of Symposia in Pure Mathematics, Vol. 19, *Combinatorics*, T.S. MOTZKIN (ed.), Amer. Math. Soc., Providence, R.I., 1971, pp. 187-204.

[18] RAY-CHAUDHURI, D.K. & R.M. WILSON, *The existence of resolvable designs*, in: *A Survey of Combinatorial Theory*, J.N. SRIVASTAVA a.o., (ed.), North-Holland/American Elsevier, Amsterdam/New York, 1973, pp. 361-376.

[19] WALLIS, W.D., A.P. STREET & J.S. WALLIS, *Combinatorics: Room squares, sum-free sets, Hadamard matrices*, Lecture Notes in Mathematics 292, Springer-Verlag, Berlin etc., 1972.

[20] WILSON, R.M., *The construction of group divisible designs and partial planes having the maximum number of lines of a given size*, in: Proceedings of the Second Chapel Hill Conference on Combinatorial Mathematics and its Applications, University of North Carolina at Chapel Hill, 1970, pp. 488-497.

[21] WILSON, R.M., *Cyclotomy and difference families in elementary abelian groups*, J. Number Theory, 4 (1972) 17-47.

[22] WILSON, R.M., *An existence theory for pairwise balanced designs, I: Composition theorems and morphisms*, J. Combinatorial Theory A, 13 (1972) 220-245.

[23] WILSON, R.M., *An existence theory for pairwise balanced designs, II: The structure of PBD-closed sets and the existence conjectures*, J. Combinatorial Theory A, 13 (1972) 246-273.

[24] WILSON, R.M., *An existence theory for pairwise balanced designs, III: Proof of the existence conjectures*, J. Combinatorial Theory, to appear.

MATHEMATICAL CENTRE TRACTS 55, 1974, 42-52

ON TRANSVERSAL DESIGNS

H. HANANI

University of the Negev, Beer Sheva, Israel

1. BASIC LEMMAS

A *design* is a pair $(X,\mathcal{B})$ where X is a finite set of *points* and $\mathcal{B}$ is a family of -not necessarily distinct- subsets B_i (called *blocks*) of X.

A *parallel class* of blocks of a design $(X,\mathcal{B})$ is a subfamily $\mathcal{B}_1 \subset \mathcal{B}$ of disjoint blocks which cover X.

In a design $(X,\mathcal{B})$ let the family $\mathcal{B}$ of blocks be composed of two subfamilies $\mathcal{B} = G \cup P$ where G is a parallel class of blocks. The elements (blocks) of G will be called *groups* and the elements (blocks) of P *-proper blocks* or for short- *blocks*. A design $(X, G \cup P)$ is a *transversal design* $T[s,\lambda,r]$ iff

(i) $|G_i| = r$ for every $G_i \in G$,
(ii) $|G| = s$,
(iii) $|G_i \cap B_j| = 1$ for every $G_i \in G$ and every $B_j \in P$,
(iv) every pairset $\{x,y : x \in G_i, y \in G_j, G_i \neq G_j\}$ is contained in exactly λ blocks of P.

It follows immediately that in $T[s,\lambda,r]$, $|X| = sr$, $|B_j| = s$ for every $B_j \in P$, and $|P| = r^2\lambda$.

Let us denote by $T(s,\lambda)$ the set of integers r for which designs $T[s,\lambda,r]$ exist. The following lemmas are evident.

LEMMA 1. $T(s,\lambda) \subset T(s',\lambda)$ *for every* $s' \leq s$.

LEMMA 2. $T(s,\lambda) \subset T(s,n\lambda)$ *for every positive integer* n.

Lemma 2 may be generalized as follows.

LEMMA 3. *If* $r \in T(s,\lambda)$ *and* $r \in T(s,\lambda')$, *then* $r \in T(s,n\lambda+n'\lambda')$ *for all non-negative integers* n *and* n'.

The following lemma has been proved by MACNEISH [5] for $\lambda = \lambda' = 1$. However, this, more general wording, does not involve any change in the proof.

LEMMA 4. *If* $r \in T(s,\lambda)$ *and* $r' \in T(s,\lambda')$, *then* $rr' \in T(s,\lambda\lambda')$.

We shall now prove

LEMMA 5. *In a transversal design* $T[s,\lambda,r]$, $s \leq (r^2\lambda-1)/(r-1)$ *holds.*

PROOF. Let $X = I_r \times I_s$. We may assume that one of the blocks is $\{(0;\sigma) : \sigma \in I_s\}$. There are exactly $r\lambda-1$ additional blocks containing the point $(0;0)$. Each of the points $(0;\sigma')$, $\sigma'=1,2,\ldots,s-1$ occurs in these blocks exactly $\lambda-1$ times and accordingly the total number of points having as first index 0 in those $r\lambda-1$ blocks is $r\lambda-1+(s-1)(\lambda-1)$. The total number of pairs $\{(0;\sigma),(0;\sigma') : \sigma,\sigma' \in I_s, \sigma \neq \sigma'\}$ occurring in the blocks is minimized if each of the blocks has the same number of points having first index 0, namely $1+(s-1)(\lambda-1)/(r\lambda-1)$ and then the total number of the said pairs is $\frac{1}{2}(r\lambda-1)[(s-1)(\lambda-1)/(r\lambda-1)+1][(s-1)(\lambda-1)/(r\lambda-1)]$.

There are exactly $(r-1)r\lambda$ blocks not containing the point $(0;0)$. The number of points with first index 0 in those blocks is $(r-1)\lambda(s-1)$ and the total number of pairs of such points is minimized if in each block there is an equal number, i.e. $(s-1)/r$, of points with first index 0.

Summing up we have for the total number P of pairs of points with first index 0

$$P = \tfrac{1}{2}s(s-1)\lambda \geq$$

$$\geq \tfrac{1}{2}s(s-1) + \tfrac{1}{2}(r\lambda-1)\left[\frac{(s-1)(\lambda-1)}{r\lambda-1} + 1\right]\frac{(s-1)(\lambda-1)}{r\lambda-1} +$$

$$+ \tfrac{1}{2}(r-1)r\lambda\frac{s-1}{r}\left[\frac{s-1}{r} - 1\right],$$

which gives $s \leq (r^2\lambda-1)/(r-1)$. □

Transversal designs satisfying $s = (r^2\lambda-1)/(r-1)$ will be called *complete transversal designs*. Transversal designs with $s < (r^2\lambda-1)/(r-1)$ will be called *incomplete transversal designs*.

If the blocks of a transversal design $T[s,\lambda,r]$ can be partitioned into $r\lambda$ parallel classes of blocks then the design will be called a *resolvable transversal design* $RT[s,\lambda,r]$. As usual the set of integers r for which designs $RT[s,\lambda,r]$ exist will be denoted by $RT(s,\lambda)$.

By adding an additional group to a resolvable transversal design $RT[s,\lambda,r]$ and adjoining each element of this group to λ distinct parallel classes of blocks we obtain

LEMMA 6. $RT(s,\lambda) \subset T(s+1,\lambda)$.

For $\lambda = 1$ the stronger result is known

LEMMA 7. $RT(s,1) = T(s+1,1)$.

Transversal designs are also known as *Orthogonal arrays*. A full description of such arrays including bibliography may be found in the book of DAMARAJU RAGHAVARAO [8, p.9-31].

2. COMPLETE TRANSVERSAL DESIGNS

2.1. Hadamard matrices

The complete transversal designs with $r = 2$, namely the designs $T[4\lambda-1,\lambda,2]$, are equivalent to the Hadamard matrices. To see this, write the groups of the design in any fixed order and in each group denote one point by +1 and the other by -1. Further write the blocks of the design as rows of a matrix and add a column of +1's. The obtained matrix is a $4\lambda \times 4\lambda$ Hadamard matrix.

The designs $T[4\lambda-1,\lambda,2]$ with $\lambda=2,3,\ldots,8$ are given herewith.

$2 \in T(7,2)$ $\qquad X = I_2 \times Z_7$

$\{(0;0),(0;1),(0;2),(0;3),(0;4),(0;5),(0;6)\}$

$\{(0;1),(0;2),(0;4),(1;0),(1;3),(1;5),(1;6)\} \bmod (-;7)$

$2 \in T(11,3)$ $\qquad X = I_2 \times Z_{11}$

$\{(0;\zeta) : \zeta \in Z_{11}\}$

$\{(1;0),(0;2^{2\alpha}),(1;2^{2\alpha+1}) : \alpha=0,1,2,3,4\} \bmod (-;11)$

$2 \in T(15,4)$ $\qquad X = I_2 \times Z_{15}$

$\{(0;\zeta) : \zeta \in Z_{15}\}$

$\{(0;1),(0;2),(0;3),(0;5),(0;6),(0;9),(0;11),(1;0),(1;4),(1;7),(1;8),(1;10),$
$(1;12),(1;13),(1;14)\} \bmod (-;15)$

$2 \in T(19,5)$ $\qquad X = I_2 \times Z_{19}$

$\{(0;\zeta) : \zeta \in Z_{19}\}$

$\{(1;0),(0;2^{2\alpha}),(1;2^{2\alpha+1}) : \alpha=0,1,\ldots,8\} \bmod (-;19)$

$2 \in T(23,6)$ $\quad X = I_2 \times Z_{23}$

$\{(0;\zeta) : \zeta \in Z_{23}\}$

$\{(1;0),(0;5^{2\alpha}),(1;5^{2\alpha+1}) : \alpha=0,1,\ldots,10\} \bmod (-;23)$

$2 \in T(27,7)$ $\quad X = I_2 \times GF(27)$ $\quad x^3 = x + 2$

$\{(0;\zeta) : \zeta \in GF(27)\}$

$\{(1;0),(0;x^{2\alpha}),(1;x^{2\alpha+1}) : \alpha=0,1,\ldots,12\} \bmod (-;27)$

$2 \in T(31,8)$ $\quad X = I_2 \times Z_{31}$

$\{(0;\zeta) : \zeta \in Z_{31}\}$

$\{(1;0),(0;3^{2\alpha}),(1;3^{2\alpha+1}) : \alpha=0,1,\ldots,14\} \bmod (-;31)$.

2.2. Projective planes

The complete transversal designs with $\lambda = 1$, namely the designs $T[q+1,1,q]$ are equivalent to the finite projective planes $PG(2,q)$. As lines in the plane may serve the blocks as well as the groups with an additional point (∞) adjoint.

It is known that finite projective planes exist whenever q is a power of a prime and accordingly we have

LEMMA 8. *If* q *is a power of a prime, then* $q \in T(q+1,1)$.

2.3. Projective geometries

It is known that finite projective geometries $PG(d,q)$ of any dimension d exist if q is a power of a prime. Such geometry enables us to construct a complete transversal design $T[(q^d-1)/(q-1),q^{d-2},q]$. To this end fix any point A of $PG(d,q)$ and define as groups all the $(q^d-1)/(q-1)$ lines through A, with A deleted. Every (d-1)-dimensional hyperplane not incident with A intersects every group in exactly one point and we may define those (d-1)-dimensional hyperplanes as blocks of the design. Through every two points of distinct groups goes exactly one line of the geometry. This line is contained in $(q^{d-1}-1)/(q-1)$ (d-1)-dimensional hyperplanes; out of these hyperplanes $(q^{d-2}-1)/(q-1)$ are incident with A, the other q^{d-2} are blocks of the design. Accordingly every pair of points in distinct groups is contained in exactly q^{d-2} blocks. Consequently we obtain

THEOREM 1. *If* q *is a power of a prime, then* $q \in T((q^d-1)/(q-1), q^{d-2})$ *for every integer* d > 1.

Two designs of the described form are given herewith.

$3 \in T(13,3)$ $\quad X = (Z_2 \cup \{\infty\}) \times Z_{13}$

$\{(\infty;\zeta) : \zeta \in Z_{13}\}$

$\{(\infty;0),(\infty;2^{4\alpha+1}),(0;2^{4\alpha}),(1;2^{4\alpha+2}),(1;2^{4\alpha+3}) : \alpha=0,1,2\}$ mod $(2;13)$

$5 \in T(31,5)$ $\quad X = (Z_4 \cup \{\infty\}) \times Z_{31}$

$\{(\infty;\zeta) : \zeta \in Z_{31}\}$

$\{(\infty;3^{10\alpha}),(\infty;3^{10\alpha+3}),(0;0),(0;3^{10\alpha+1}),(0;3^{10\alpha+4}),(1;3^{10\alpha+5}),(1;3^{10\alpha+7}),$
$(1;3^{10\alpha+9}),(2;3^{10\alpha+8}),(3;3^{10\alpha+2}),(3;3^{10\alpha+6}) : \alpha=0,1,2\}$ mod $(4;31)$.

3. INCOMPLETE TRANSVERSAL DESIGNS

3.1. Affine geometries

In the case $\lambda = 1$ an incomplete resolvable transversal design RT[q,1,q] representing an affine plane AG(2,q) is obtained from a complete transversal design T[q+1,1,q] representing a projective plane PG(2,q) by omitting one group. Such a design exists whenever q is a power of a prime.

In the general case, finite affine geometries AG(d,q) of any dimension d exist if q is a power of a prime. Such geometry enables us to construct a resolvable transversal design $RT[q^{d-1},q^{d-2},q]$ as follows:
Take any class of parallel lines of the geometry as groups and every (d-1)-dimensional hyperplane which does not contain a group, as a block. In this construction we obtain q^{d-1} classes of q parallel blocks each. Consequently

THEOREM 2. *If* q *is a power of a prime then for every integer* d,
$q \in RT(q^{d-1},q^{d-2})$.

A design of the described form is given herewith

$3 \in T(9,3)$ $\quad X = Z_3 \times GF(9)$ $\quad x^2 = 2x + 1$

$\{(0;0),(1;x^{2\alpha}),(2;x^{2\alpha+1}) : \alpha=0,1,2,3\}$ mod $(3;9)$

3.2. Latin squares

It has been observed long ago, that the existence of a transversal design $T[s,1,r]$ is equivalent to the existence of $s-2$ mutually orthogonal Latin squares of order r. As the Latin squares have become more popular, most of the results are stated in a form convenient in that field. We shall word these results in a way accepted for transversal designs.

From lemmas 4 (with $\lambda = \lambda' = 1$) and 8 follows the result of MACNEISH [6].

LEMMA 9. *If* $r = \Pi p_i^{\alpha_i}$ *is the factorization of* r *into powers of distinct primes, then* $r \in T(s,1)$, *where* $s-1 = \min p_i^{\alpha_i}$.

Some other most important results in this field are listed below.

(1) $r \in T(s,1)$ whenever $r > (3s)^{91}$ (CHOWLA, ERDŐS & STRAUSS [3]).
This result has been improved by ROGERS [9] and lately by WILSON who proved

(2) for sufficient large s, $r \in T(s,1)$ whenever $r > s^{17}$ (WILSON [10]).
Further it has been proved

(3) for every $r > 6$, $r \in T(4,1)$ holds (BOSE, PARKER and SHRIKHANDE [7,2,1]),

(4) for every $r > 51$, $r \in T(5,1)$ holds (HANANI [5]),

(5) for every $r > 62$, $r \in T(7,1)$ holds (HANANI [5]),

(6) for every $r > 90$, $r \in T(8,1)$ holds (WILSON [10]).

For further reference we mention a result by DULMAGE, JOHNSON & MENDELSOHN [4] who proved the existence of 5 mutually orthogonal Latin squares of order 12, or, in our notation

LEMMA 10. $12 \in T(7,1)$.

3.3. General transversal designs

Let a transversal design $T[\sigma,\lambda,\rho]$ exist. Delete some elements from one of the groups so that in this group $\rho' \leq \rho$ elements are left. The blocks will be of size σ and $\sigma-1$. It follows immediately

LEMMA 11. *If* $\rho \in T(\sigma,\lambda)$ *and* $\rho' \leq \rho$, *and if for a given* s, $\{\sigma,\sigma-1\} \subset RT(s,1)$ *and* $\{\rho,\rho'\} \subset T(s,\lambda)$, *then* $r = \rho(s-1) + \rho' \in T(s,\lambda)$.

We shall now prove

THEOREM 3. *For every* $r \geq 1$ *and every* $\lambda > 1$, $r \in T(7,\lambda)$ *holds.*

PROOF. For r = 1 the lemma is trivial. Further, it follows from lemmas 4 and 11 that in order to prove our lemma for every r > 1 it is sufficient to prove it for the factors of 60. This is done presently.

For r = 2 we proved in section 2.1 that $2 \in T(7,2)$ and $2 \in T(11,3)$. By lemmas 1 and 3 it follows

(3.1) $\qquad 2 \in T(7,\lambda)$ for every $\lambda > 1$.

For r = 3 it follows from theorem 1 (with d = 3) that $3 \in T(13,3)$. Further we have

$3 \in RT(6,2)$ $\qquad X = Z_3 \times (Z_3 \times I_2)$

$\{(0;0,0),(0;1,0),(1;2,0),(2;0,1),(2;1,1),(1;2,1)\}$ mod $(3;3,-)$

$\{(0;2\alpha,0),(\alpha;2\alpha+1,0),(2\alpha;2\alpha+2,0),(0;\alpha,1),(\alpha;\alpha+1,1),(2\alpha;\alpha+2,1)\}$ mod $(3;-,-)$, $\alpha=0,1,2$.

Consequently by lemmas 6, 1 and 3 we have

(3.2) $\qquad 3 \in T(7,\lambda)$ for every $\lambda > 1$.

For r = 4 we have

$4 \in RT(8,2)$ $\qquad X = GF(4) \times (Z_7 \cup \{\infty\})$ $\qquad x^2 = x + 1$

$\{(0;\zeta) : \zeta \in (Z_7 \cup \{\infty\})\}$ mod $(4;-)$

$\{(0;\infty),(0;0),(x^0;3^\alpha),(x^1;3^{\alpha+2}),(x^2;3^{\alpha+4}) : \alpha=0,1\}$ mod $(4;7)$

$4 \in RT(8,3)$ $\qquad X = GF(4) \times I_8$ $\qquad x^2 = x + 1$

$\{(0;0),(0;1),(x^\alpha;2),(x^{\alpha+1};3),(0;4),(x^{\alpha+1};5),(x^\alpha;6),(x^\alpha;7)\}$ mod $(4;-)$, $\alpha=0,1,2$
$\{(0;0),(x^\alpha;1),(0;2),(x^{\alpha+1};3),(x^{\alpha+1};4),(0;5),(x^\alpha;6),(x^{\alpha+1};7)\}$ mod $(4;-)$, $\alpha=0,1,2$
$\{(0;0),(x^\alpha;1),(x^{\alpha+1};2),(0;3),(x^{\alpha+1};4),(x^\alpha;5),(0;6),(x^\alpha;7)\}$ mod $(4;-)$, $\alpha=0,1,2$
$\{(0;0),(x^{\alpha+1};1),(x^{\alpha+1};2),(x^{\alpha+1};3),(x^\alpha;4),(x^\alpha;5),(x^\alpha;6),(0;7)\}$ mod $(4;-)$, $\alpha=0,1,2$.

Considering lemmas 1, 3 and 6 we have

(3.3) $\qquad 4 \in RT(8,\lambda)$ and $4 \in T(9,\lambda)$ for every $\lambda > 1$.

For $r = 5$

$5 \in RT(10,2)$ $\quad X = Z_5 \times (Z_5 \times I_2)$

$\{(0;0,\alpha),(4;2^{2\alpha},0),(1;2^{2\alpha+1},0),(2;2^{2\alpha},1),(3;2^{2\alpha+1},1) : \alpha=0,1\}$ mod $(5;5,-)$

$\{(0;0,\alpha),(2;2^{2\alpha},0),(3;2^{2\alpha+1},0),(3;2^{2\alpha},1),(2;2^{2\alpha+1},1) : \alpha=0,1\}$ mod $(5;5,-)$.

$5 \in RT(7,3)$ $\quad X = Z_5 \times Z_7$

$\{(0;\zeta) : \zeta \in Z_7\}$ mod $(5;-)$

$\{(0;0),(2^{\beta};3^{3\alpha}),(2^{\beta+1};3^{3\alpha+2}),(2^{\beta+3};3^{3\alpha+1}) : \alpha=0,1\}$ mod $(5;7)$, $\beta=0,1$.

Consequently by lemmas 1, 3 and 6

(3.4) $\quad 5 \in RT(7,\lambda)$ and $5 \in T(8,\lambda)$ for every $\lambda > 1$.

For $r = 6$

$6 \in T(7,2)$ $\quad X = (Z_5 \cup \{\infty\}) \times Z_7$

$\{(\infty;\zeta) : \zeta \in Z_7\}$ 2 times

$\{(\infty;0),(0;3^{3\alpha}),(2^{\beta};3^{3\alpha+2}),(2^{\beta+2};3^{3\alpha+1}) : \alpha=0,1\}$ mod $(5;7)$, $\beta=0,1$.

$6 \in T(7,3)$ $\quad X = (Z_5 \cup \{\infty\}) \times Z_7$

$\{(\infty;\zeta) : \zeta \in Z_7\}$ 3 times

$\{(\infty;0),(0;2),(0;3),(0;5),(2^{2\alpha};4),(2^{2\alpha+1};1),(2^{2\alpha+3};6)\}$ mod $(5;7)$, $\alpha=0,1$

$\{(\infty;0),(0;3^{3\alpha}),(1;3^{3\alpha+2}),(4;3^{3\alpha+1}) : \alpha=0,1\}$ mod $(5;7)$.

Consequently, by lemma 3, we have

(3.5) $\quad 6 \in T(7,\lambda)$ for every $\lambda > 1$.

For $r = 10$

$10 \in T(7,2)$ $\quad X = (Z_9 \cup \{\infty\}) \times Z_7$

$\{(\infty;\zeta) : \zeta \in Z_7\}$ 2 times

$\{(0;\zeta) : \zeta \in Z_7\}$ mod $(9;-)$

$\{(\infty;0),(0;3^{3\alpha}),(1;3^{3\alpha+2}),(2;3^{3\alpha+4}),(4;3^{3\alpha+5}),(6;3^{3\alpha+1}),(7;3^{3\alpha+3})\}$ mod $(9;7)$, $\alpha=0,1$

$\{(0;0),(1;3^{3\alpha+2}),(2;3^{3\alpha}),(7;3^{3\alpha+1}) : \alpha=0,1\}$ mod $(9;7)$.

$10 \in T(8,3)$ $\quad X = (Z_3 \times Z_3 \cup \{\infty\}) \times (Z_2 \times Z_2 \times Z_2)$

$\{(\infty;\zeta) : \zeta \in Z_2 \times Z_2 \times Z_2\}$ 3 times

$\{(0,0;\zeta) : \zeta \in Z_2 \times Z_2 \times Z_2\}$ mod $(3,3;-)$

$\{(\infty;0,0,0),(0,0;0,0,1),(0,0;0,1,0),(0,1;0,1,1),(0,2;1,1,0),(1,2;1,0,1),$
$(1,2;1,0,0),(2,0;1,1,1)\}$ mod $(3,3;2,2,2)$

$\{(\infty;0,0,0),(0,0;0,1,0),(0,0;1,0,0),(0,1;1,1,0),(0,2;1,0,1),(1,1;0,1,1),$
$(1,1;0,0,1),(2,1;1,1,1)\}$ mod $(3,3;2,2,2)$

$\{(\infty;0,0,0),(0,0;1,0,0),(0,0;0,0,1),(0,1;1,0,1),(0,2;0,1,1),(1,0;1,1,0),$
$(1,0;0,1,0),(2,2;1,1,1)\}$ mod $(3,3;2,2,2)$

$\{(0,0;0,1,1),(0,1;1,0,1),(0,2;1,1,0),(1,0;0,1,0),(1,1;1,0,0),(1,2;0,0,1),$
$(2,0;0,0,0),(2,0;1,1,1)\}$ mod $(3,3;2,2,2)$.

By lemmas 1 and 3 it follows

(3.6) $\quad 10 \in T(7,\lambda)$ for every $\lambda > 1$.

For $r = 15$

$15 \in RT(7,2)$ $\quad X = (Z_3 \times Z_5) \times Z_7$

$\{(0,0;\zeta) : \zeta \in Z_7\}$ mod $(3,5;-)$, 2 times

$\{(0,2^{\gamma+3};0),(2^{\alpha+\beta},0;3^{3\alpha}),(2^{\alpha+\beta},2^{\gamma};3^{3\alpha+2}),(2^{\alpha+\beta},2^{\gamma+2};3^{3\alpha+4}) : \alpha=0,1\}$
mod $(3,5;7)$, $\beta=0,1$, $\gamma=0,1$.

$15 \in RT(7,3)$ $\quad X = Z_{15} \times Z_7$

$\{(0;\zeta) : \zeta \in Z_7\}$ mod $(15;-)$, 3 times

$\{(0;0),(1;3^{\alpha}),(2;3^{\alpha+1}),(4;3^{\alpha+2}),(5;3^{\alpha+3}),(8;3^{\alpha+4}),(10;3^{\alpha+5})\}$ mod $(15;7)$,
$\alpha=0,1,\ldots,5$.

By lemmas 3 and 6 we have

(3.7) $\quad 15 \in RT(7,\lambda)$ and $15 \in T(8,\lambda)$ for every $\lambda > 1$.

For $r = 20$

$20 \in T(7,2)$ $\quad X = (Z_{19} \cup \{\infty\}) \times Z_7$

$\{(\infty;\zeta) : \zeta \in Z_7\}$ 2 times

$\{(\infty;0),(2^{6\alpha+\mu};3^{2\alpha}),(2^{6\alpha+\mu};3^{2\alpha+3}) : \alpha=0,1,2\}$ mod $(19;7)$, $\mu=2,12$

$\{(0;0),(2^{6\alpha+9\beta};3^{2\alpha+4\gamma}),(2^{6\alpha+9\beta+1};3^{2\alpha+4\gamma+3}) : \alpha=0,1,2\}$ mod $(19;7)$, $\beta=0,1$, $\gamma=0,1$.

$20 \in T(7,3)$ $\quad X = (Z_{19} \cup \{\infty\}) \times Z_7$

$\{(\infty;\zeta) : \zeta \in Z_7\}$ 3 times

$\{(\infty;0),(2^{6\alpha+\mu};3^{2\alpha}),(2^{6\alpha+\mu};3^{2\alpha+3}) : \alpha=0,1,2\}$ mod $(19;7)$, $\mu=6,14,16$

$\{(0;0),(2^{6\alpha+2\beta+9\gamma};3^{2\alpha}),(2^{6\alpha+2\beta+9\gamma+1};3^{2\alpha+3}) : \alpha=0,1,2\}$ mod $(19;7)$, $\beta=0,1,2$, $\gamma=0,1$.

By lemma 3 we have

(3.8) $\quad 20 \in T(7,\lambda)$ for every $\lambda > 1$.

For $r = 12$ see lemma 10 and for $r = 30$ and $r = 60$ apply lemma 11: in the case of $r = 30$ with parameters $\sigma = 8$, $\rho = 4$ and $\rho' = 2$, and in the case $r = 60$ with parameters $\sigma = 8$, $\rho = 8$ and $\rho' = 4$. □

REFERENCES

[1] BOSE, R.C., E.T. PARKER & S.S. SHRIKHANDE, *Further results on the construction of mutually orthogonal Latin squares and the falsity of Euler's conjecture*, Canad. J. Math., 12 (1960) 189-203.

[2] BOSE, R.C. & S.S. SHRIKHANDE, *On the construction of sets of mutually orthogonal Latin squares and the falsity of a conjecture of Euler*, Trans. Amer. Math. Soc., 95 (1960) 191-209.

[3] CHOWLA, S., P. ERDŐS & E.G. STRAUSS, *On the maximal number of pairwise orthogonal Latin squares of a given order*, Canad. J. Math., 12 (1960) 204-208.

[4] DULMAGE, A.L., D.M. JOHNSON & N.S. MENDELSOHN, *Orthomorphisms of groups and orthogonal Latin squares, I*, Canad. J. Math., 13 (1961) 356-372.

[5] HANANI, H., *On the number of orthogonal Latin squares*, J. Combinatorial Theory, 8 (1970) 247-271.

[6] MACNEISH, H.F., *Euler squares*, Ann. of Math., 23 (1922) 221-227.

[7] PARKER, E., *Construction of some sets of mutually orthogonal Latin squares*, Proc. Amer. Math. Soc., 10 (1959) 946-949.

[8] RAGHAVARAO, D., *Constructions and combinatorial problems in design of experiments*, J. Wiley & Sons, New York, 1971.

[9] ROGERS, K., *A note on orthogonal Latin squares*, Pacific J. Math., 14 (1964) 1395-1397.

[10] WILSON, R.M., *Concerning the number of mutually orthogonal Latin squares*, in print.

FINITE GEOMETRY

MATHEMATICAL CENTRE TRACTS 55, 1974, 55-63

COMBINATORICS OF FINITE GEOMETRIES

A. BARLOTTI

Università di Bologna, Bologna, Italy

In this lecture we intend to present a brief survey of very recent results. We shall be interested in the development of some topics considered in section 3.2 of DEMBOWSKI's book [21] (Combinatorics of finite planes) and in problems connected with the existence of finite geometrical structures.

1. THE STUDY OF SYSTEMS AXIOMATIZING FINITE PLANES AND DESIGNS*)

1.1. If $P = (\mathcal{p},\mathcal{L},I)$ is a projective plane of order n, then the following properties hold:

(1) $|\mathcal{p}| = n^2 + n + 1$;

(2) $|\mathcal{L}| = n^2 + n + 1$;

(3) $[p] = n + 1$, for every $p \in \mathcal{p}$;

(4) $[L] = n + 1$, for every $L \in \mathcal{L}$;

(5) $[p,q] = 1$, for $p,q \in \mathcal{p}$ and $p \neq q$;

(6) $[L,M] = 1$, for $L,M \in \mathcal{L}$ and $L \neq M$.

Conversely, if in a structure P consisting of a set $\mathcal{p}$ of points, a set $\mathcal{L}$ of lines with an incidence defined in $\mathcal{p} \times \mathcal{L}$, properties (1) to (6) hold, with $n \geq 2$, then P is a (non-degenerate) projective plane. Moreover, the above properties (1) to (6) are a redundant set of conditions to ensure that P is a projective plane. A subset of (1) - (6) is called a *complete* system when its properties are sufficient to imply that P is a projective plane. A complete system is called *minimal* if none of its proper subsets is also complete.

---

*) See also DEMBOWSKI [21], pp. 138-139. Here, and in what follows we shall use the symbols used in this book.

In M. HALL [30] it is proved that (2), (4), (5) and its "dual"*) (1), (3), (6) are complete systems. All complete minimal systems of (1) to (6) were found in [1] and the following list was given:

(1) (3) (6) , (2) (4) (5),
(1) (4) (5) , (2) (3) (6),
(1) (2) (3) (5), (1) (2) (4) (6).

If $p \cup L$ is not empty, the systems

(3) (4) (5) (3) (4) (6)

are complete and minimal.

Further, if $|p| \geq 2$ the system (3) (5) (6) is also complete and minimal, and the dual of this statement also holds.

G. CORSI [19] considered a refinement of this problem using, instead of properties (1),...,(6) the twelve properties obtained by replacing property (i) by the properties (i'), in which " = " is replaced by " ≤ ", and (i"), in which " = " is replaced by " ≥ ".**) The full list of complete minimal subsystems of (1'),...,(6") is given in DEMBOWSKI [21] pp. 138-139. Notice, however, that in the statement of the result given there we must replace the words "nondegenerate incidence structure" by "incidence structure containing at least two points or two lines".

Clearly, if we add some further hypotheses to (1') - (6") the list of complete minimal subsystems of these will be modified. We list here some results in this direction.

a) A. BASILE [4] added the hypothesis that the structure P is *nondegenerate*.

b) C. BERNASCONI [5] studied the problem for a *connected* and nondegenerate structure P.

The fact that P is connected adds to the list of complete minimal systems given in [4] the following:

(1") (3') (4') (5"), (2") (3') (4') (6"),
(1") (3') (5), (2") (4') (6),

and requires the cancellation of four systems which are no longer minimal, viz.:

---

*) Clearly, properties (2),(4),(6) are respectively the duals of (1),(3),(5).

**) Logical questions connected with the replacement of condition (i) by (i') or (i") are considered in R. MAGARI [34].

(1") (3') (4) (5"), (2") (3) (4') (6"),
(1") (3') (4') (5") (6"), (2") (3') (4') (5") (6").

c) M. CRISMALE [20] found all the complete minimal subsets of the system obtained from (1') - (6") by replacing (3'), (3"), (4'), (4") respectively by the following properties (3'), (3"), (4'), (4"):

$(\hat{3}')$ $b_1 \leq n+1$, $(\hat{3}")$ $b_1 \geq n+1$,
$(\hat{4}')$ $v_1 \leq n+1$, $(\hat{4}")$ $v_1 \geq n+1$,

where b_1 is the average number of lines on a point and v_1 is the average number of points on a line. The list of the complete minimal subsets is the following:

(1") (2') $(\hat{3}")$ (6'), (1') (2") $(\hat{4}")$ (5'),
(1") (2') $(\hat{3}")$ (5'), (1') (2") $(\hat{4}")$ (6'),
(1") $(\hat{3})$ $(\hat{4}")$ (5") (6'), (2") $(\hat{3}")$ $(\hat{4})$ (5') (6"),
(1") $(\hat{3})$ $(\hat{4}")$ (5), (2") $(\hat{3}")$ $(\hat{4})$ (6),
(1) $(\hat{3}")$ $(\hat{4}")$ (6'), (2) $(\hat{3}")$ $(\hat{4}")$ (5'),
(1) $(\hat{3}")$ $(\hat{4}")$ (5'), (2) $(\hat{3}")$ $(\hat{4}")$ (6'),
(1") $(\hat{3}")$ $(\hat{4})$ (6), (2") $(\hat{3})$ $(\hat{4}")$ (5),
(1") $(\hat{3}")$ $(\hat{4})$ (5') (6"), (2") $(\hat{3})$ $(\hat{4}")$ (5") (6').

1.2. Similar questions can be studied for affine planes considering together with properties (3) and (5) above the following:

(7) $|p| = n^2$,
(8) $|L| = n(n+1)$,
(9) $[L] = n$ for every $L \in \mathcal{L}$.

Partial results, needed as lemmas to obtain other results, were obtained by OSTROM [37] and DEMBOWSKI & OSTROM [22]. The systematic search for the complete minimal subsystems was done by U. OLIVERI [36] for the set (3) (5) (6') (7) (8) (9) and by P. BRUTTI [15] for (3') (3") (5') (5") (7') (7") (8') (8") (9') (9") with the additional condition that the structure is nondegenerate. C. BERNASCONI [5] studied how the hypothesis that the structure is connected modifies the list given in [15].

1.3. Similar questions may be studied for other finite structures; this was done for inversive planes by R. BUMCROT & D. KNEE [17] and for projective designs by C. BERNASCONI [6]. We shall give here the results obtained in [6].

Let D be a structure consisting of a set $\mathcal{P}$ of points, a set $\mathcal{B}$ of blocks and an incidence defined in $\mathcal{P} \times \mathcal{B}$ for which the following properties hold:

(D1) $\quad |\mathcal{P}| = v = 1 + \frac{k(k-1)}{\lambda}$,

(D2) $\quad |\mathcal{B}| = b = 1 + \frac{k(k-1)}{\lambda}$,

(D3) $\quad [p] = r = k$ for every $p \in \mathcal{P}$,

(D4) $\quad [B] = k$ for every $B \in \mathcal{B}$,

(D5) $\quad [pp'] = \lambda$ for $p,p' \in \mathcal{P}$ and $p \neq p'$,

(D6) $\quad [BB'] = \mu = \lambda$ for $B,B' \in \mathcal{B}$ and $B \neq B'$,

with the hypothesis $\lambda > 0$ and $k - \lambda \geq 2$. The above list is the extension of (1) to (6) of 1.1 to the case of projective designs, and in [6] it is proved that the list of complete minimal systems of (D1) to (D6) is the same (and with the same additional conditions) as the list given in 1.1 for the case of projective planes. This result shows that projective designs can be characterized by cardinality conditions, without any symmetry assumptions.

In connection with this fact we wish to observe that D.A. DRAKE [29] has proved that in a finite affine Hjelmslev plane, cardinality assumptions cannot replace the parallel axiom.

Other papers in connection with the above topics are [7], [40] and [41]. The following theorem (N.G. DE BRUIJN & P. ERDÖS [14]) represents a seminal result in the study of the problems considered in the first section:

Let there be given n points $a_1,\ldots,a_n$ and denote by $A_1,\ldots,A_m$ blocks of points such that we have:

(i) $|A_i| \geq 2$;

(ii) each pair (a_i,a_j) is contained in one and only one block.

Then we have $m \geq n$, with equality occurring only if either the system is of the type:

$A_1 = (a_1,a_2,\ldots,a_{n-1})$, $A_2 = (a_1,a_n)$, $A_3 = (a_2,a_n),\ldots,$ $A_n = (a_{n-1},a_n)$

or if the system is a projective plane, with the lines given by $A_1,\ldots,A_n$.

2. REFERENCES TO RECENT RESULTS IN OTHER TOPICS

Representation of nets and planes by sets of mutually orthogonal Latin squares is one of the central topics in combinatorics of finite geometric structures. For recent results and problems in this field we refer to A. BARLOTTI [3], G. DÉNES & A.D. KEEDWELL [23] and P. HOHLER [33].

The purpose of Galois geometries is to study geometry of finite spaces. Non-linear properties are of particular interest. Characterizations of algebraic varieties in finite spaces are illustrated in G. TALLINI [43]. Problems related to (k;n)-arcs and (k;n)-caps of given "kind" are considered in M. TALLINI SCAFATI [44].

In connection with the "packing problem" for k-caps, we quote the following results:

$m(4,3) = 20$ due to G. PELLEGRINO [38],
$m(5,3) = 56$ due to R. HILL [31].

We wish also to point out how a representation of the plane in higher dimensional space may sometimes simplify the study of the arcs in the plane. (See R.C. BOSE [8].)

Combinatorial arguments may be of big help in problems connected with the existence or non-existence and characterization of finite geometric structures. See R.H. BRUCK [10], A. BRUEN [11], A. BRUEN & J.C. FISHER [12], J. COFMAN [18] and R.H.F. DENNISTON [24 to 28].

REFERENCES

[1] BARLOTTI, A., *Un'osservazione sulle proprietà che caratterizzano un piano grafico finito*, Boll. Un. Mat. Ital., 17 (1962) 394-398.

[2] BARLOTTI, A., *Some classical and modern topics in finite geometrical structures*, in: *A survey of combinatorial theory*, J.N. SRIVASTAVA a.o. (eds.), North-Holland Publ. Cy., Amsterdam, 1973.

[3] BARLOTTI, A., *Alcune questioni combinatorie nello studio delle strutture geometriche*, in: Atti Convegno Teorie Combinatorie, Acc. Lincei, Rome 1973, to appear.

[4] BASILE, A., *Sugli insiemi di proprietà che definiscono un piano grafico finito*, Le Matematiche, 25 (1970) 84-95.

[5] BERNASCONI, C., *Strutture di incidenza connesse e definizione assiomatica di piani grafici e affini*, Ann. Univ. Ferrara, to appear.

[6] BERNASCONI, C., *Sistemi di assiomi che caratterizzano i disegni proiettivi*, to appear.

[7] BISCARINI, P., *Sets of axioms for finite inversive planes*, to appear.

[8] BOSE, R.C., *On a representation of Hughes planes*, in: Proc. Internat. Conf. on Projective Planes, M.J. KALLAHER & T.G. OSTROM (eds.), Washington State Univ. Press, 1973, pp.27-57.

[9] BRAMWELL, D.L. & B.J. WILSON, *The (11,3)-arcs of the Galois plane of order 5*, Proc. Cambridge Philos. Soc., 74 (1973) 247-250.

[10] BRUCK, R.H., *Construction problems in finite projective spaces*, in: *Finite geometric structures and their applications*, C.I.M.E. II ciclo 1972, Ed. Cremonese, Rome, 1973, pp.105-188.

[11] BRUEN, A., *Blocking sets in finite projective planes*, SIAM J. Appl. Math., 21 (1971) 380-392.

[12] BRUEN, A. & J.C. FISHER, *Arcs and ovals in derivable planes*, Math. Z., 125 (1972) 122-128.

[13] BRUEN, A. & J.C. FISHER, *Blocking-sets, k-arcs and nets of order ten*, Advances in Math., 10 (1973) 317-320.

[14] BRUIJN, N.G. DE & P. ERDÖS, *On a combinatorial problem*, Kon. Nederl. Akad. Wetensch. Proc. A, 51 (1948) 1277-1279 (= Indag. Math., 10 (1948) 421-423).

[15] BRUTTI, P., *Sistemi di assiomi che definiscono un piano affine di ordine n*, Ann. Univ. Ferrara Sez VII, 14 (1969) 109-118.

[16] BUEKENHOUT, F. & R. METZ, *On circular spaces having no disjoint circles*, to appear.

[17] BUMCROT, R. & D. KNEE, private communication.

[18] COFMAN, J., *On combinatorics of finite projective spaces*, in: Proc. Internat. Conf. on Projective Planes, M.J. KALLAHER & T.G. OSTROM (eds.), Washington State Univ. Press, 1973, pp.59-70.

[19] CORSI, G., *Sui sistemi minimi di assiomi atti a definite un piano grafico finito*, Rendic. Sem. Mat. Padova, 34 (1964) 160-175.

[20] CRISMALE, M., *Sui sistemi minimi di assiomi atti a definite un piano proiettivo finito*, to appear.

[21] DEMBOWSKI, P., *Finite geometries*, Ergebnisse der Mathematik 44, Springer-Verlag, Berlin etc., 1968.

[22] DEMBOWSKI, P. & T.G. OSTROM, *Planes of order n with collineation groups of order n^2*, Math. Z., 103 (1968) 239-258.

[23] DÉNES, J. & A.D. KEEDWELL, *Latin squares and their applications*, Acad. Press, New York and London, 1974.

[24] DENNISTON, R.H.F., *Some packings of projective spaces*, Rend. Acc. Naz. Lincei (8), 52 (1972) 36-40.

[25] DENNISTON, R.H.F., *Cyclic packings of the projective space of order 8*, Rend. Acc. Naz. Lincei (8), 54 (1973) 373-377.

[26] DENNISTON, R.H.F., *Packings of PG (3,q)*, in: *Finite geometric structures and their applications*, C.I.M.E. II ciclo 1972, Ed. Cremonese, Rome, 1973, pp.193-199.

[27] DENNISTON, R.H.F., *Spreads which are not subregular*, Glasnik Mat. Ser. III, 8 (1973) 3-5.

[28] DENNISTON, R.H.F., *Some spreads which contain reguli without being subregular*, to appear.

[29] DRAKE, D.A., *Near affine Hjelmslev planes*, J. Comb. Theory, 16 (1974) 34-50.

[30] HALL, M. Jr., *The theory of groups*, Mac Millan, New York, 1959.

[31] HILL, R., *On the largest size of cap in $S_{5,3}$*, Rend. Acc. Naz. Lincei, to appear.

[32] HILL, R., *Caps and groups*, to appear.

[33] HOHLER, P., *Eigenschaften von vollständigen Systemen orthogonaler Lateinischer Quadrate, die bestimmte affine Ebenen repräsentieren*, J. of Geometry 2 (1972) 161-174.

[34] MAGARI, R., *Sui sistemi di assiomi "minimali" per una data teoria*, Boll. Un. Mat. Ital., 19 (1964) 423-435.

[35] MENICHETTI, G., *q-archi completi nei piani di Hall di ordine* $q = 2^k$, to appear.

[36] OLIVERI, U., *Alcune proprietà che caratterizzano un piano affino finito*, Le Matematiche, 22 (1967) 397-402.

[37] OSTROM, T.G., *Semi translation planes*, Trans. Amer. Math. Soc., 111 (1964) 1-18.

[38] PELLEGRINO, G., *Sul massimo ordine delle calotte in* $S_{4,q}$, Le Matematiche, 25 (1970) 149-157.

[39] PELLEGRINO, G., *Procedimenti geometrici per la construzione di alcune classi di calotte complete in* $S_{r,3}$, Boll. Un. Mat. Ital. (4), 5 (1972) 109-115.

[40] REIMAN, I., *Su una proprietà dei piani grafici finiti*, Rend. Acc. Naz. Lincei, 35 (1963) 279-281.

[41] REIMAN, I., *Su una proprietà dei due disegni*, Rend. Mat. e Appl., 1 (1968) 75-81.

[42] SEGRE, B., *Proprietà elementari relative ai segmenti ed alle coniche sopra un campo qualsiasi ed una congettura di Seppa Ilkka per il caso dei campi di Galois*, Ann. Mat. Pura Appl. (4), 96 (1973) 289-337.

[43] TALLINI, G., *Graphic characterization of algebraic varieties in a Galois space*, in: Atti Convegno Teorie Combinatorie, Acc. Lincei, Rome 1973, to appear.

[44] TALLINI SCAFATI, M., *The k-sets of type (m,n) in a Galois space* $S_{r,q}$ *(r≥2)*, in: Atti Convegno Teorie Combinatorie, Acc. Lincei, Rome 1973, to appear.

[45] THAS, J.A., *Connection between the n-dimensional affine space* $A_{n,q}$ *and the curve C, with equation* $y = x^q$, *of the affine plane* A_{2,q^n}, Rend. Trieste, 2 (1970) 146-151.

[46] THAS, J.A., *A combinatorial problem*, Geometriae Dedicata, 1 (1973) 236-240.

[47] THAS, J.A., *4-gonal configurations*, in: *Finite geometric structures and their applications*, C.I.M.E. II ciclo 1972, Ed. Cremonese, Rome, 1973, pp.249-263.

[48] THAS, J.A., *On 4-gonal configurations*, Geometriae Dedicata, 2 (1973) 317-326.

[49] THAS, J.A., *Flocks of finite egglike inversive planes*, in: *Finite geometric structures and their applications*, C.I.M.E. II ciclo 1972, Ed. Cremonese, Rome, 1973, pp.189-191.

[50] THAS, J.A., *Some results concerning $\{(q+1)(n-1);n\}$-arcs and $\{(q+1)(n-1)+1;n\}$-arcs in finite projective planes of order q*, to appear.

[51] THAS, J.A., *On 4-gonal configurations with parameters $r = q^2+1$ and $k = q+1$*, to appear.

MATHEMATICAL CENTRE TRACTS 55, 1974, 64-112

ON FINITE NON-COMMUTATIVE AFFINE SPACES

J. ANDRÉ

Universität des Saarlandes, D 66 Saarbrücken, GFR

INTRODUCTION

We consider incidence structures (cf. DEMBOWSKI [10,p.1]) consisting of a non-void set of elements called *points*, certain subsets of points called *lines*, an operation sometimes called *join*, mapping every ordered pair of different points surjectively onto the set of all lines and finally a binary relation on the lines called *parallelism*, all these things with additional properties gained in a natural way. The most known structures of this type are the well-known affine spaces (e.g. in the sense of TAMASCHKE [20,21]) but also many other examples have been developed in recent times (see e.g. DEMBOWSKI [10], PICKERT [18], SPERNER [19], ARNOLD [6], WILLE [26], also ANDRÉ [1], BACHMANN [8]).

In very general spaces with a *commutative* join many results are known in the meantime, especially by the extensive work of WILLE [26]. For spaces with a non-commutative join, however, almost nothing is known up to now. It is the purpose of this paper to develop some results just for such "non-commutative spaces". But we do not want to consider too general spaces of such a type. For this reason we introduce some further axioms in such a way that the space under consideration becomes an affine space (in the usual sense) if additionally the join is commutative, thus obtaining "non-commutative affine spaces". If we do so in a suitable way we get the so-called *quasi-affine spaces* and some of their particular types, especially the *near-affine spaces (fastaffine Räume)*.

As we shall see in this note these geometries can be applied to the theory of groups, esp. permutation groups (see also ANDRÉ [3]), to algebra, esp. nearfields (see also ANDRÉ [4], BACHMANN [8]), to the theory of combinatorial designs (see also ANDRÉ [2]) and to the foundations of geometry (see also ANDRÉ [5] and BACHMANN [8]).

There may possibly be other applications. For instance, by considering the set of all countable sequences with elements of a Kalscheuer-nearfield (cf. KALSCHEUER [16]) and with a convergent series of the squares

of their absolute values we get a generalization of Hilbert space which could give new aspects on functional analysis, perhaps also new deep insights in quantum mechanics and generalizations.*)

In the first chapter the basic definitions, esp. of the quasi-affine and nearaffine spaces, will be stated (cf. sections 1 and 3). A simple but, as I suppose, far reaching application to transitive permutation groups (not necessarily of finite degree) is contributed in section 2. In the last section of this chapter axioms of configurations, some of them analogous to that of DESARGUES in affine spaces, are described; their significance for geometric structures will be given in the last chapter. It may also be possible to generalize e.g. the Reidemeister-, Thomsen- and Klingenberg-configurations (see e.g. KLINGENBERG [17], HUGHES & PIPER [12], PICKERT [18]) to nearaffine spaces and then to characterize such spaces in some other ways (e.g. algebraically). All results stated in this first chapter are valid for both finite and infinite spaces.

For the second chapter, however, the condition of finiteness for all spaces under consideration is essential. We introduce there the concepts of order, dimension, subspaces and others for finite nearaffine spaces in a natural way. We gain many theorems concerning their detailed structure some of them being analogous to well-known properties in affine spaces. One of the main results is that two different points are incident with either exactly one or exactly n lines where n is the order (i.e. the number of points on a line) of the space under consideration (cf. chapter II, theorem 6.1). This leads to another type of join introduced in definition 6.1. It is commutative but deeper properties of it are not yet known.

In the last chapter we consider finite nearaffine spaces with further properties which, roughly spoken, assume the existence of "sufficiently many" points, lines etc. with special properties. Under such conditions one can prove the little Desargues condition (cf. chapter III, section 1) and as a consequence the existence of sufficiently many translations. Nothing, however, is known to the author which hypotheses imply the general Desargues theorem. But if this is true then the nearaffine space can be described as a space over a nearfield, even in the infinite case; this will be sketched in the last section.

---

*) For a "geometrical" theory of quantum mechanics see e.g. VARADARAJAN [22,23], with further references. For perhaps possible generalizations in this direction see e.g. JORDAN, VON NEUMANN & WIGNER [15], JORDAN [13] or (with respect to non-commutative lattices) JORDAN [14]. (See also problem (7) of the appendix.)

Finally we shall list some unsolved problems related to quasi-affine and nearaffine spaces whose solutions might possibly lead to a deeper understanding of those structures and their relations to other fields of mathematics.

I. BASIC CONCEPTS

1. QUASI-AFFINE SPACES

We consider structures of the type

$$S = (X, L, \sqcup, \|) \tag{1.1}$$

with the following *basic hypotheses*:

(1) X is an arbitrary non-void set whose elements are called *points*. They are denoted by small latin letters.

(2) L is a subset of the powerset $P(X)$ of X; their elements are called *lines* (*Linien*). They are denoted by capital latin letters.

(3) $\sqcup$ is a surjective mapping of the set of all ordered pairs (x,y) of different points onto L:

$$\begin{cases} \sqcup: X^2 \setminus \Delta_X \longrightarrow\!\!\!\!\!\rightarrow L, \\ \\ (x,y) \longmapsto x \sqcup y, \qquad (x \neq y) . \end{cases} \tag{1.2}$$

The line $x \sqcup y$ is called the *connection-line* or *join from* x *to* y (*in this order*).*)

(4) $\|$ is a binary equivalence relation on L called *parallelism*.

In this way S becomes an incidence structure (in the sense of DEMBOWSKI [10] **)) where ϵ is the incidence relation between X and L.

We shall give some further conditions on S in such a way that S becomes an affine space (in the usual sense) if we additionally assume the commutativity of $\sqcup$. We subdivide these axioms into four classes: (L) *conditions on lines*, (P) *conditions on parallelism*, (R) *a condition of rich-*

---

*) It is sometimes useful to define also $x \sqcup x =: \{x\}$.

**) There the lines are called *blocks*.

ness, and (T) *a condition on similar triangles* (see also ANDRÉ [2]).

DEFINITION 1.1. A structure $S = (X,L,\sqcup,\|)$ with the basic hypotheses (1) to (4) is called a *quasi-affine space* if the following conditions hold:

(L1) $x,y \in x\sqcup y$ *for all* $x,y \in X$ *with* $x \neq y$.

(L2) $z \in (x\sqcup y)\setminus\{x\} \iff x\sqcup y = x\sqcup z$.

(L3) $x\sqcup y = y\sqcup x = x\sqcup z \Rightarrow x\sqcup z = z\sqcup x$.

(P1) *For* $L \in \mathsf{L}$ *and* $x \in X$ *there exists exactly one line* $L' \| L$ *such that* $L' = x\sqcup y$ *for a suitable* $y \in X$ (Euclid's axiom of parallelism). This line is denoted by

$$(1.3) \qquad (x \| L) .$$

(P2) *If* $G = x\sqcup y = y\sqcup x$ *and* $L \| G$ *then* $x',y' \in L$ *with* $x' \neq y'$ *imply* $x'\sqcup y' = y'\sqcup x'$.

(R) *There exist at least two lines in* S.

(T) *If* x,y,z *are pairwise different and* x',y' *are different points with* $x\sqcup y \| x'\sqcup y'$ *then*

$$(1.4) \qquad (x' \| x\sqcup z) \cap (y' \| y\sqcup z) \neq \emptyset. \ ^{*)}$$

Consequences and remarks

(a) Condition (L1) implies that every line contains at least two points.

(b) A point x on a line L such that $L = x\sqcup y$ for a suitable $y \in X$ is called a *base point* (*Aufpunkt*) of L.

A simple consequence of (L3) is

(c) The following conditions are equivalent:
  (α) L has at least two base points.
  (β) Every point of L is a base point of L.
  (γ) If x,y are different points on L then $L = x\sqcup y = y\sqcup x$.

(d) A line with a condition given in (c) is called a *straight line* (*Gerade*). The set of all straight lines of the space S is denoted by G. Lines not being straight are called *proper lines*.

(e) Condition (P2) means that every line parallel to a straight line is also straight.

---

*) See TAMASCHKE [21,p.319] for an analogous condition on affine spaces.

(f) If we specialize (T) by x = x' we get the following affine version of a *Veblen-condition*:

(V) *Let* x,y,z *be pairwise different points,* y' $\in$ x$\sqcup$y, *then*

(1.4') $\qquad (x\sqcup z) \cap (y' \;||\; y\sqcup z) \neq \emptyset.$

Note that this intersection [as well as (1.4)] need not consist of only one point.

The following theorem is easy to prove (see also ANDRÉ [2,theorem 2.1]).

THEOREM 1.1. *A quasi-affine space* $S = (X,L,\sqcup,||)$ *with a commutative join* $\sqcup$ *is an affine space (of dimension* ≥ 2*).*

DEFINITION 1.2. Let $A \subseteq X$. Two points $x,y \in A$ are called *joinable (verbindbar) with respect to* A, denoted by

$$x \underset{A}{\sim} y,$$

if either x = y or there exist finitely many points $x = x_0,x_1,\ldots,x_n = y$ such that all $x_i \sqcup x_{i+1}$ ($i\in\{0,1,\ldots,n-1\}$) are straight lines totally contained in A (i.e. x and y can be connected by a finite chain of straight lines all completely belonging to A).

It is straightforward that $\underset{A}{\sim}$ is an equivalence relation on X. For A = X we simply say that x and y are *joinable* and simplify the sign $\underset{A}{\sim}$ to ~. The equivalence classes with respect to ~ all reduce to single points iff S contains no straight lines.

2. EXAMPLES RELATED TO PERMUTATION GROUPS

Let X be a non-void set (not necessarily finite) and Γ a permutation group acting on X. For $x \in X$ let Γ_x be, as usual, the *stabilizer*, i.e. the subgroup of all those $\gamma \in \Gamma$ leaving x fixed. Moreover, assume always that $x \neq y$ implies $\Gamma_x \neq \Gamma_y$. We define a *join* $\sqcup$ on X by

(2.1) $\qquad x\sqcup y := \{x\} \cup \Gamma_x(y),$

thus obtaining a set L of lines on X. Such a line with a base point x therefore consists of x and an orbit of Γ_x different from x. A parallelism on L is defined by

(2.2) $\quad L \parallel L' :\Longleftrightarrow L' = \gamma(L)$ for a suitable $\gamma \in \Gamma$.

In this way we get a structure $I_\Gamma = (X, \mathcal{L}, \sqcup, \parallel)$ called the *group space* of Γ possessing all basic hypotheses [cf. section 1, (1) to (4)]. Moreover, the following property is easy to verify.

PROPOSITION 2.1. *If Γ is a transitive but not doubly transitive permutation group on X then the space I_Γ defined by (2.1) and (2.2) is a quasi-affine space. Moreover, for all $\gamma \in \Gamma$ we have*

(2.3) $\quad \gamma(x \sqcup y) = \gamma(x) \sqcup \gamma(y)$

for all $x,y \in X$, thus γ being an automorphism of I_Γ.

We give some characterizations for additional conditions on Γ or I_Γ.

PROPOSITION 2.2. *The condition*

(2.4) $\quad x \sqcup y \parallel y \sqcup x, \qquad (x,y \in X,\ x \neq y)$

holds iff Γ contains a γ exchanging x and y. *)

PROOF. $\gamma(x) = y$, $\gamma(y) = x$ and (2.3) imply $x \sqcup y \parallel \gamma(x \sqcup y) = \gamma(x) \sqcup \gamma(y) = y \sqcup x$, hence (2.4).

Conversely assume now (2.4). By (2.2) there exists a $\delta \in \Gamma$ with $y \sqcup x = \delta(x \sqcup y) = \delta(x) \sqcup \delta(y)$. Using (2.1) we have $\delta(x) = y$ and $\delta(y) \in \Gamma_y(x)$. This yields the existence of an $\eta \in \Gamma_y$ mapping $\delta(y)$ on x. The permutation $\gamma := \eta\delta \in \Gamma$ exchanges x and y.

PROPOSITION 2.3. *The equation*

(2.5) $\quad x \sqcup y = y \sqcup x, \qquad (x,y \in X,\ x \neq y),$

holds iff for any $\gamma \in \Gamma_x \backslash \Gamma_y$ there exists a $\gamma' \in \Gamma_y \backslash \Gamma_x$ such that

---

*) Another interesting characterization of this exchanging condition is given in WIELANDT [25,theorem 16.4,p.45]. Moreover, it is easy to prove by similar methods that the orbit $\neq \{x\}$ of Γ_x given by $(x \parallel y \sqcup x)$ is the reflection (cf. [25,§16,p.44]) of the orbit determined by $x \sqcup y$, independently of the special choice of y. Other relations between Γ and I_Γ will be published elsewhere.

$$\gamma(y) = \gamma'(x)$$

and a corresponding relation holds if the roles of x *and* y *are interchanged.*

PROOF. See ANDRÉ [3,Satz 1.1]. □

THEOREM 2.1. *Let* Γ *be a transitive but not doubly transitive permutation group on* X *such that for all different points* x, y *and any* $\gamma \in \Gamma_x \backslash \Gamma_y$ *there exists a* $\gamma' \in \Gamma_y \backslash \Gamma_x$ *with* $\gamma(y) = \gamma'(x)$. *Then* Γ *is isomorphic (as permutation group) to the group of all dilatations (i.e. translations or homotheties) of a desarguesian affine space (of dimension* ≥ 2*).*

This is a consequence of theorem 1.1 and proposition 2.3. For further details cf. ANDRÉ [3].

PROPOSITION 2.4. *The equivalence classes on* X *with respect to the joinable relation* ~ (cf. definition 1.2) *form a complete block system of* Γ (see e.g. WIELANDT [25,p.12]), *i.e.*

(1) *for any equivalence class* X_i *and any* $\gamma \in \Gamma$ *we have either* $\gamma(X_i) = X_i$ *or* $X_i \cap \gamma(X_i) = \emptyset$, *and*

(2) *for any two classes* X_i *and* X_j *there exists a* $\gamma \in \Gamma$ *with* $X_j = \gamma(X_i)$.

This follows because the image $\gamma(x \sqcup y)$ [$\gamma \in \Gamma$] of a straight line is again straight.

THEOREM 2.2. *A transitive group* Γ *whose group space* I_Γ *contains at least one straight line is either doubly transitive or imprimitive.*

PROOF. Assume Γ to be transitive but not doubly transitive. Let G be a straight line of I_Γ then $\emptyset \neq G \subset X$. Let $x \in X \backslash G$ and $y \in X \backslash \{x\}$ such that $x \sqcup y \| G$. Then $\gamma(x) \in G$ implies $\gamma(y) \in G$ for all $\gamma \in \Gamma$. Due to Rudio's theorem (see e.g. WIELANDT [25, theorem 8.1]) Γ is imprimitive. □

For considerations occurring in the next section and in chapter III the following concept may be useful. Again let Γ be a group operating transitively but not doubly transitively on X.

DEFINITION 2.1. A $\tau \in \Gamma$ is called a *translation* if either $\tau = 1$ (the identity) or τ acts fixpointfree on X and

(2.6) $\qquad x \sqcup \tau(x) \;\|\; y \sqcup \tau(y)$

holds for all $x,y \in X$.

The set T of all translations need not be a subgroup of Γ but trivially $\tau \in T$ and $\gamma \in \Gamma$ imply $\gamma\tau\gamma^{-1} \in T$, and if additionally (2.4) holds for all $x,y \in X$ then $\tau \in T$ implies $\tau^{-1} \in T$ due to

$$x \sqcup \tau^{-1}(x) =$$
$$= \tau^{-1}(\tau(x) \sqcup x) \;\|\; \tau(x) \sqcup x \;\|\; x \sqcup \tau(x) \;\|\; y \sqcup \tau(y) \;\|\; \tau(y) \sqcup y \;\|\; \tau^{-1}(\tau(y) \sqcup y) =$$
$$= y \sqcup \tau^{-1}(y).$$

PROPOSITION 2.5. *A fixpointfree permutation $\tau \in \Gamma$ belongs to T iff for all $x,y \in X$ there exists a $\gamma \in \Gamma$ such that*

(2.7) $\qquad y = \gamma(x) \quad \textit{and} \quad y = \tau^{-1}\gamma\tau(x)$

hold.

PROOF. Assume (2.7); then

$$y \sqcup \tau(y) = \gamma(x) \sqcup \tau\tau^{-1}\gamma\tau(x) = \gamma(x \sqcup \tau(x)) \;\|\; x \sqcup \tau(x).$$

Conversely suppose (2.6) for τ. According to the transitivity of Γ there exists a $\delta \in \Gamma$ with $\delta(x) = y$, Now (2.6) and (P1) yield

$$y \sqcup \delta\tau(x) = \delta(x \sqcup \tau(x)) = y \sqcup \tau(y),$$

hence $\tau(y) \in \Gamma_y \delta\tau(x)$ due to (2.1). Thus we have

$$y \in \tau^{-1}\Gamma_y \delta\tau(x),$$

so that we can find an $\eta \in \Gamma_y$ with $y = \tau^{-1}\eta\delta\tau(x)$. The permutation $\gamma := \eta\delta \in \Gamma$ has all properties of (2.7). □

DEFINITION 2.2. A translation τ is called *straight* (*strikt*) if either $\tau = 1$ or $x \sqcup \tau(x)$ is straight.

REMARK. Because of (P2) this condition does not depend on the special choice of x.

The straight translations will play an important role in some special type of quasi-affine spaces considered in the next section. It should be important to characterize those groups Γ whose translations form a subgroup, especially a (sharply) transitive subgroup. It is an interesting but up to now unsolved problem to characterize the spaces I_Γ in a purely geometric way. This problem is only solved for some very special types of permutation groups Γ. (See also problem (2) of the appendix in this paper.)

3. NEARAFFINE SPACES

In this section we specialize the concept of a quasi-affine space in such a way that we essentially assume the existence of "sufficiently many" straight lines.

DEFINITION 3.1. A quasi-affine space $F = (X, L, \sqcup, ||)$ is called a *nearaffine space* (*fastaffiner Raum*) if the following additional conditions hold in F (they are subdivided into two classes: one *condition concerning parallelism* (P) and two *conditions* (G) *concerning straight lines*):

(P3) *For all* $x,y \in X$ *with* $x \neq y$ *we have* *)

$$x\sqcup y \;||\; y\sqcup x .$$

(G1) *Every line* L *meets every straight line* $G \neq L$ *in at most one point, i.e.* $|G \cap L| \leq 1$.

(G2) Chain condition. *All points of* X *are joinable in the sense of definition 1.2.*

As a consequence of (T) [cf. definition 1.1] and (P3) the following condition holds.

THEOREM 3.1. Condition for closed parallelograms (Pa). *Let* x,y,z *be pairwise different points with* $x\sqcup y \neq x\sqcup z$; *then*

$$(3.1) \qquad (z \;||\; x\sqcup y) \cap (y \;||\; x\sqcup z) \neq \emptyset$$

---

*) This is exactly the condition (2.4).

(but the intersection need not consist of only one point).

PROOF. Consider the diagonal y⌊z. Using (P3) we have x⌊z || z⌊x , x⌊y || y⌊x and y⌊z || z⌊y . If we apply (T) on the triangle x,y,z and on z,y we obtain (3.1). □

REMARK. For nearaffine spaces (V) and (Pa) are consequences of (T). It is an unsolved problem, even in the finite case, whether conversely (T) follows from (V) and (Pa). For the analysis of finite nearaffine spaces as done in the following chapters, however, we *only* need the conditions (V) and (Pa).

A fundamental property of nearaffine spaces is given by the following theorem.

THEOREM 3.2. *All lines of a nearaffine space have the same number of points.*

PROOF. The proof will be given in several steps: First using (V) and (G1) we show that two intersecting straight lines have the same number of points. By the use of (G2) we secondly prove that any two straight lines have equally many points. The third step is to verify that two lines with a common base point have the same number of points. For that we heavily use (G2) in the form of an induction on the number of straight lines connecting two points of both lines, different from their common base point. The complete theorem is then a simple consequence of this result. For further details cf. ANDRÉ [2,I,theorem 3.1]. □

Note that for the proof of this theorem we use the axioms (G) and (V) but not (Pa) and instead of (P3) only the weaker condition x⌊y || x'⌊y' ⇒ y⌊x || y'⌊x' (by using prop. 3.3).

DEFINITION 3.2. The (common) number of points on a line (which may of course be infinite) is called the *order* of the nearaffine space. It is usually denoted by n.

REMARKS.

(1) By (L1) we always have n ≥ 2. For n = 2 we get x⌊y = {x,y} = y⌊x, i.e. an affine space. Hence n ≥ 3 holds for any proper nearaffine space.

(2) It is well-known that the numbers of points of the orbits of Γ_x different from x differ in general (see e.g. WIELANDT [25,chapter III]). Therefore theorem 3.2 does not hold in a general quasi-affine space.

EXAMPLES OF NEARAFFINE SPACES. We consider the following permutation group Γ acting on the set X. Let I be an index set consisting of at least two elements. Moreover, let F be a set in which a multiplication and for every $i \in I$ an addition $+_i$ is defined in such a way that $(F,+_i,\cdot)$ becomes a (not necessarily planar) nearfield (with the distributive law $\alpha(\beta +_i \gamma) = \alpha\beta +_i \alpha\gamma$ for all $\alpha,\beta,\gamma \in F$) *) and all these nearfields have the same neutral element 0 with respect to their additions $+_i$. Let X be defined by

$$X := \{(\xi_i)_{i\in I} \mid \xi_i \in F,\ \xi_i \neq 0 \text{ for only finitely many } i \in I\}. \tag{3.2}$$

Define addition and multiplication by

$$(\xi_i)_{i\in I} + (\eta_i)_{i\in I} := (\xi_i +_i \eta_i)_{i\in I} \tag{3.3}$$

and

$$\alpha(\xi_i)_{i\in I} := (\alpha\xi_i)_{i\in I} \tag{3.4}$$

resp., $(\alpha,\xi_i,\eta_i \in F)$. Then (X,+) is an abelian group and every $\alpha \in F\backslash\{0\}$ becomes an automorphism $x \mapsto \alpha x$ of (X,+). Moreover, $\alpha x = \beta x$ implies $\alpha = \beta$ or $x = 0$.

The group Γ of all permutations

$$x \mapsto \alpha x+v, \qquad (x,v \in X,\ \alpha \in F\backslash\{0\}) \tag{3.5}$$

is transitive but (due to $|I| \geq 2$) not doubly transitive on X. Their translations (cf. definition 2.1) are exactly the mappings (3.5) with $\alpha = 1$; they form a sharply transitive normal subgroup of Γ. A translation $x \mapsto x+v$ is straight (cf. definition 2.2) iff v belongs to the so-called *quasi-nucleus* (*Quasikern*) Q of X defined by

$$Q := \{v \in X \mid \forall\alpha,\beta \in F\ \ \exists\gamma \in F \text{ with } \alpha v+\beta v = \gamma v\}. \tag{3.6}$$

The lines $x \sqcup y$ of I_Γ can also be described by

---

*) For the definition and properties of nearfields see e.g. DEMBOWSKI [10,p.33] (with the left distributive law instead of the right one quoted here).

(3.7) $\qquad x\sqcup y := F(y-x) + x := \{\alpha(y-x) + x \mid \alpha \in F\}.$

Such a line is straight iff $y-x \in Q$. The parallelism of two lines is given by

(3.8) $\qquad Fz + v \parallel Fz' + v' :\Longleftrightarrow Fz = Fz', \qquad (z,z' \in X\backslash\{0\},\ v,v' \in X).$

The cardinality $|I|$ of the index set I is sometimes called the *dimension* of $\mathcal{I}_\Gamma$; it is denoted by

(3.9) $\qquad \mathrm{Dim}\ \mathcal{I}_\Gamma := |I|$

and does only depend on the geometrical structure of $\mathcal{I}_\Gamma$.

DEFINITION 3.3. The spaces $\mathcal{I}_\Gamma$ defined by the group (3.5) are sometimes called *nearfield spaces*.

Due to proposition 2.1 all nearfield spaces are quasi-affine. But we can prove more.

PROPOSITION 3.1. *A nearfield space is nearaffine.*

The proofs of (P3) and (G2) are straightforward. For the proof of (G1) we essentially have to use that

$$\alpha v + \beta x = \alpha' v + \beta' x, \qquad (v \in Q\backslash\{0\},\ x \in X\backslash Fv)$$

implies $\alpha = \alpha'$ and $\beta = \beta'$. For the proof of this proposition cf. ANDRÉ [4, Satz 4.17].*) Also the following proposition is easy to verify.

PROPOSITION 3.2. *Let $\mathcal{I}_\Gamma$ be a nearfield space. Then the additions $+_i$ given in* (3.3) *do not depend on* i *(so that we can put $+_i =: +$) iff the following additional condition holds in $\mathcal{I}_\Gamma$.*

(G3) Triangle condition. *For any two different straight lines* G *and* G' *with the common point* x *there exist* $y \in G\backslash\{x\}$ *and* $y' \in G'\backslash\{x\}$ *such that* $y\sqcup y'$ *is straight.*

*) In that paper one can find an algebraic theory of the structures defined by (3.2) to (3.4), there called *Fastvektorräume* (*nearvectorspaces*). The geometric consequences will be given in ANDRÉ [5].

We conclude this section with a proposition valid for general near-affine planes which will be useful in the next chapter.

PROPOSITION 3.3. *Let* x *and* y *be different points on a straight line* G *and let* L *be a line different from* G *with* x *as a base point. Then*

$L \cap (y \parallel L) = \emptyset$.

PROOF. Assume $z \in L \cap (y \parallel L)$. Then $L = x \sqcup z$ and $(y \parallel L) = y \sqcup z$ by (L2). Using (P3) we get

$$z \sqcup x \parallel L \parallel (y \parallel L) \parallel z \sqcup y ,$$

by (P1) thus $z \sqcup x = z \sqcup y$, hence $x,y \in G \cap (z \sqcup x)$ contradicting (G1) because G is straight. □

REMARK. It is easy to see that for the proof of proposition 3.3 we only need the condition $x \sqcup y \parallel x' \sqcup y' \Rightarrow y \sqcup x \parallel y' \sqcup x'$ instead of the stronger one (P3).

4. CONFIGURATIONS

In order to get deeper properties of nearaffine spaces it is often necessary to add further hypotheses. They may be of group-theoretical nature, e.g. by assuming the existence of "sufficiently many" collineations (automorphisms) with certain properties. On the other hand it is sometimes useful to make hypotheses of geometrical nature, i.e. by requiring geometrical configurations analogous to the Desargues or Pappos configurations in affine spaces. We first formulate three types of Desargues conditions.

(D1) Little Desargues configuration. *If* $x,x',y,y',z,z' \in X$ *are pairwise different,* $x \sqcup x' \parallel y \sqcup y' \parallel z \sqcup z'$ *are straight and pairwise different, then* $x \sqcup y \parallel x' \sqcup y'$ *and* $x \sqcup z \parallel x' \sqcup z'$ *imply* $y \sqcup z \parallel y' \sqcup z'$.

(D2) Desargues configuration (first kind). *If* $u,x,x',y,y',z,z' \in X$ *are pairwise different,* $u \sqcup x$ *is straight and* $u \sqcup y$, $u \sqcup z$ *are lines different from each other and from* $u \sqcup x$, *then* $x' \in u \sqcup x$, $y' \in u \sqcup y$, $z' \in u \sqcup z$, $x \sqcup y \parallel x' \sqcup y'$ *and* $x \sqcup z \parallel x' \sqcup z'$ *imply* $y \sqcup z \parallel y' \sqcup z'$.

(D3) Desargues configuration (second kind). *If* $u,x,x',y,y',z,z' \in X$ *are pairwise different,* $u \sqcup x$, $u \sqcup y$ *and* $u \sqcup z$ *are pairwise different lines, and if* $x \sqcup y \parallel x' \sqcup y'$ *and* $x \sqcup z \parallel x' \sqcup z'$ *are straight then* $y \sqcup z \parallel y' \sqcup z'$.

The following configuration trivially holds in any affine space.

(Di) Diagonal condition for quadrilaterals. *If* x,y,z,w ∈ X *are pairwise different such that* x⌊y, y⌊z, z⌊w, w⌊x *and the one diagonal* x⌊z *are straight then also the other diagonal* y⌊w *is straight.*

Nearaffine spaces with (D1), (D2) and (Di) are called *desarguesian.**) The following theorem is fundamental.

THEOREM 4.1. *Any nearfield space* (cf. definition 3.3) *is desarguesian.*

For the proof see ANDRÉ [5]. Conversely all desarguesian nearaffine spaces are essentially of that type; for more details see section 4 of chapter III and ANDRÉ [5].

There exist non-desarguesian nearaffine spaces of arbitrarily high dimension. They can be constructed from affine spaces over the reals by "refracting" lines in a suitable way, cf. ANDRÉ [2,chapter I]. In chapter III of this paper we shall see that in all finite nearaffine spaces,being no planes,the little Desargues condition (D1) holds. In the infinite case this problem remains open.

It is also possible to state a condition analogous to Pappos' configuration (ANDRÉ [2,5]). A nearfield space is pappian iff the multiplication of the nearfields $(F,+_i,\cdot)$ is commutative, hence all those nearfields become (commutative) fields. If additionally (G3) holds then by proposition 3.2 the set X forms a vectorspace over a field and $\mathcal{F}$ becomes an affine pappian space.

II. GENERAL THEORY OF FINITE NEARAFFINE SPACES

In this and the following chapter $\mathcal{F} := (X,L,\sqcup,||)$ is always a fixed *finite* nearaffine space (see chapter I, definition 3.1), i.e. X is a finite set. Hence the order n of $\mathcal{F}$ (cf. chapter I, definition 3.2) is a natural number. The case n = 2 is trivial (cf. the first remark after definition 3.2 in chapter I), hence *we always assume* n ≥ 3. Instead of (T) we only suppose (V) and (Pa) (cf. chapter I, sections 1 and 3, esp. the remark after theorem 3.1).

---

*) (D3) is then a conclusion of these conditions as will be shown in ANDRÉ [5]. But it is unknown whether (D2) is a consequence of (D1) and (D3), or whether (D1) follows from (D2) and (D3).

1. SUBSPACES

DEFINITION 1.1. A subset U of points of a nearaffine space F is called an *affine subspace* or simply *subspace* of F if the following two conditions hold.

(S1) $x,y \in U$, $x \neq y$ *imply* $x \sqcup y \subseteq U$.

(S2) *Any two points of* U *are joinable with respect to* U (cf. chapter I, definition 1.2).

Trivially $\emptyset$, the single points, the straight lines and the whole set X are all subspaces. The property of U being a subspace will be denoted by

$$U \leq X. \tag{1.1}$$

PROPOSITION 1.1. *Let* U *be a subspace and* G *a straight line. Then either* $G \leq U$ *or* $|U \cap G| \leq 1$.

PROPOSITION 1.2. *Let* U *be a subspace,* $x \in U$ *and* L *a line totally contained in* U, *then also* $(x \parallel L) \subseteq U$. (For the definition of $(x \parallel L)$ see chapter I, (1.3).)

PROOF. Let y be a base point (cf. chapter I,section 1,(b)) of L. If $x = y$ or $L = x \sqcup y$ then $(x \parallel L) = L \subseteq U$. Assume therefore $x \neq y$ and $L \neq x \sqcup y$. We first suppose in addition that $x \sqcup y$ is straight. Due to $n \geq 3$ there exists a $z \in (x \sqcup y) \setminus \{x,y\}$, obviously $z \notin L$. Let x' be any point of $(x \parallel L) \setminus \{x\}$, then $x' \notin x \sqcup y$. By the Veblen condition (V) there exists a $y' \in (z \sqcup x') \cap L \subseteq U$. This gives $x' \in U$, hence $(x \parallel L) \subseteq U$ if $x \sqcup y$ is straight. To complete the proof we apply (S2) for an induction on the number of straight lines in U connecting x with y. □

COROLLARY. *If* U *is a subspace of* F *then all axioms of a nearaffine space, possibly except* (R) (cf. chapter I,sections 1 and 3) *hold in the space*

$$F_U := (U, L_U, \sqcup_{|U}, \parallel_{|U}), \tag{1.2}$$

where $L_U := \{L \in L \mid L \subseteq U\}$ *and* $\sqcup_{|U}$, $\parallel_{|U}$ *are the restrictions of* $\sqcup$, $\parallel$ *resp., to* U *(in the usual sense). If* (G3) *holds in* F *then also in* F_U.

PROPOSITION 1.3. *Let* U *be a subspace and* G *a straight line with* G ∩ U ≠ ∅. *Then*

$$[U,G] := \bigcup_{y\in U} (y \parallel G) \tag{1.3}$$

is also a subspace.

PROOF.

(1) If G ≤ U then [U,G] = U is a subspace. Assume now G ≰ U, i.e. $|G\cap U| = 1$ by proposition 1.1.

(2) Let z ∈ V := [U,G]. Then by (1.3) and proposition 1.1 we have $|(z \parallel G) \cap U| = 1$. Define $\bar{z}_G$ or simply $\bar{z}$, the *projection* of z on U in the direction G, by

$$\{\bar{z}_G\} := \{\bar{z}\} := (z \parallel G) \cap U. \tag{1.4}$$

(3) Let z,z' be two different points on V. If $\bar{z} = \bar{z}'$ then $z \sqcup z' = (z \parallel G) \subseteq V$. Assume now $\bar{z} \neq \bar{z}'$. Then $\bar{z} \sqcup \bar{z}' \subseteq U$ according to (S1) for U. By (Pa) (cf. chapter I, theorem 3.1) and the definition of V we obtain $(z \parallel \bar{z} \sqcup \bar{z}') \subseteq V$. Applying now (V) to this line and $z \sqcup z'$ we finally conclude $z \sqcup z' \subseteq V$. This gives (S1) for V.

(4) By (S2) for U there is a chain of straight lines in U connecting $\bar{z}$ with $\bar{z}'$. Because $z \sqcup \bar{z}$ and $z' \sqcup \bar{z}'$ are straight (or $z = \bar{z}$, or $z' = \bar{z}'$) there also exists a chain of straight lines totally contained in V connecting z with z'. This proves (S2) for V. □

COROLLARY. *If* $|U\cap G| = 1$ *then*

$$|[U,G]| = |U||G| = |U|n\ , \tag{1.5}$$

if n *is the order of the space under consideration.*

THEOREM 1.1. *The number of points of a non-void subspace of a finite affine space of order* n *is a power of* n.

PROOF. Let V be a non-void subspace. The theorem is true if $|V| = 1$, i.e. V is a point. Use induction on $|V|$ and assume $|V| > 1$. Let U be a *proper* subspace of V with a maximal number of points. Let x ∈ U and y ∈ V\U. Due to (S2) for V a chain of straight lines completely contained in V connects

x with y. Hence there is a straight line G through x contained in V but not in U. Thus $U < [U,G] \leq V$, hence $[U,G] = V$ because of the maximality of U. By (1.5) the induction hypothesis yields the theorem. □

DEFINITION 1.2. Let U be a non-void subspace. Then the exponent d in

$$|U| = n^d \tag{1.6}$$

(cf. theorem 1.1) is called the *dimension* *) of U; it is denoted by

$$d := \mathrm{Dim}\ U. \tag{1.7}$$

By convention

$$\mathrm{Dim}\ \emptyset := -1. \tag{1.7'}$$

The dimension of the whole space X of $\mathcal{F}$ is sometimes called the *dimension* Dim $\mathcal{F}$ *of* $\mathcal{F}$.

DEFINITION 1.3. A *hyperplane* of a nearaffine space $\mathcal{F}$ is a proper maximal subspace.

Hence, if H is a hyperplane then

$$\mathrm{Dim}\ H = \mathrm{Dim}\ \mathcal{F} - 1. \tag{1.8}$$

Finally we remark

$$U \leq V \leq X, \quad \mathrm{Dim}\ U = \mathrm{Dim}\ V \Rightarrow U = V. \tag{1.9}$$

2. SOME NUMBER PROPERTIES

Let $\mathcal{F}$ be a nearaffine space of order n and dimension d. Then $\mathcal{F}$ contains n^d points. We shall state further number-theorems for such spaces.

---

*) As we shall see in chapter III this concept of dimension coincides with that given in chapter I, (3.9).

PROPOSITION 2.1. *The number of lines parallel to a given straight line is* n^{d-1}.

PROOF. Let G be a fixed straight line and g the number of lines parallel to G. By (P1) and (P2) (cf. chapter I, section 1) every point is contained in exactly one line parallel to G. Hence we have $|G|g = ng = |X| = n^d$. □

PROPOSITION 2.2. *The number of lines possessing a given point as base point is*

$$\frac{n^d - 1}{n - 1} = 1 + n + \ldots + n^{d-1}.$$

PROOF. Let x be the given point. By (L2) every point $y \neq x$ contains exactly one line with x as a base point. This proves the proposition by counting arguments. □

The axioms (P1) and (P2) yield:

PROPOSITION 2.3. *The number* γ *of straight lines through a fixed point does not depend on the special choice of this point.*

REMARK. We call this number γ the *class* of the nearaffine space under consideration.

PROPOSITION 2.4. *The number of all lines incident with a given fixed point is*

$$(2.1) \qquad r := \gamma + n\left(\frac{n^d - 1}{n - 1} - \gamma\right) = n\,\frac{n^d - 1}{n - 1} - (n-1)\gamma.$$

PROOF. Let $x \in X$ be fixed. The number of straight lines through x is γ by proposition 2.3. Hence, due to proposition 2.2, the number of proper lines with x as base point is $(n^d-1)/(n-1) - \gamma$. The number of points $y \neq x$ such that $x \sqcup y$ is a proper line is

$$(n-1)\left((n^d-1)/(n-1) - \gamma\right).$$

Now (2.1) follows by the fact that proper lines with different base points are also different (by consequence (c) of chapter I, section 1). □

THEOREM 2.1. *A finite nearaffine space* F *considered as incidence structure* (with respect to points and lines and with the incidence relation ϵ, cf. DEMBOWSKI [10,p.1]) *is a tactical configuration, i.e. every line is incident with the same number* k *of points and dually every point is incident with the same number* r *of lines* (see e.g. DEMBOWSKI [10,p.4-5]).

If F *has order* n, *dimension* d *and class* γ, *if, moreover,* v *and* b *are the number of points and lines resp., then*

$$(2.2)\qquad \begin{cases} v = n^d, \quad b = n^{d-1}\gamma + n^d\left(\dfrac{n^d - 1}{n - 1} - \gamma\right), \\ k = n, \quad r = \gamma + n\left(\dfrac{n^d - 1}{n - 1} - \gamma\right). \end{cases}$$

PROOF. By theorem 3.2 (chapter I) we have k = n. By definition 1.2 (cf. also theorem 1.1) the number of points is $v = n^d$. The formula for r is (2.1), that for b follows from

$$(2.3)\qquad bk = vr$$

being valid for all tactical configurations (cf. DEMBOWSKI [10,p.5(9')]). □

PROPOSITION 2.5. *In a finite nearaffine space of order* n, *dimension* d *and class* γ *the number of all straight lines is* $n^{d-1}\gamma$.

PROOF. By the propositions 2.2 and 2.3 the number of all proper lines with a given point as base point is $(n^d-1)/(n-1) - \gamma$, hence by consequence (c) (cf. chapter I, section 1) the number of all proper lines is $n^d((n^d-1)/(n-1) - \gamma)$. Our proposition follows from theorem 2.1. □

3. A DEPENDENCE RELATION ON THE STRAIGHT LINES OF A BUNDLE

Let F be a finite nearaffine space, and L and G the set of all lines and straight lines resp. Let x be a fixed point in F. Then L_x is by definition the set of all lines with x as a base point, the *bundle* generated by x. Morover, let $G_x := G \cap L_x$.

Let $G_1,\ldots,G_k \in G_x$; then define

$$(3.1)\qquad \begin{cases} [G_1] := G_1 \\ [G_1,\ldots,G_k] := [[G_1,\ldots,G_{k-1}],G_k]. \end{cases}$$

Obviously $[G_1,\ldots,G_k]$ is a subspace. We shall prove that it does not depend on the order of the G_i's.

LEMMA 3.1. $[G_1,\ldots,G_{k-2},G_{k-1},G_k] = [G_1,\ldots,G_{k-2},G_k,G_{k-1}]$.

PROOF. Define $[G_1,\ldots,G_{k-2}] =: U$. First by (1.3) we get

$$[U,G_{k-1}] \leq [[U,G_k],G_{k-1}] = [U,G_k,G_{k-1}].$$

Moreover,

$$G_k \leq [U,G_k] \leq [U,G_k,G_{k-1}].$$

Due to proposition 1.2 we get

$$(v \parallel G_k) \leq [U,G_k,G_{k-1}]$$

for all $v \in [U,G_{k-1}]$. Hence

$$[U,G_{k-1},G_k] = \bigcup_{v\in[U,G_{k-1}]} (v \parallel G_k) \leq [U,G_k,G_{k-1}].$$

By symmetry we also have the converse inequality, thus the lemma. □

LEMMA 3.2. $[G_1,\ldots,G_{i-1},G_i,\ldots,G_k] = [G_1,\ldots,G_i,G_{i-1},\ldots,G_k]$.

PROOF. This follows from lemma 3.1 because of

$$[G_1,\ldots,G_{i-1},G_i,\ldots,G_k] = [[G_1,\ldots,G_i],\ldots,G_k]. \quad \square$$

Every permutation of $1,\ldots,k$ can be generated by transpositions of neighboured numbers. This gives

PROPOSITION 3.1. *The subspace* $[G_1,\ldots,G_k]$ *does not depend on the order of the* G_i*'s.*

DEFINITION 3.1. Let $G,G_1,\ldots,G_k \in G_x$. We call G *dependent on* $G_1,\ldots,G_k$, denoted by

$$G \text{ dep } \{G_1,\ldots,G_k\}, \tag{3.2}$$

iff $G \leq [G_1,\ldots,G_k]$.

THEOREM 3.1. *The set* G_x *together with the dependence relation given in definition 3.1 forms a dependence space* (see e.g. VAN DER WAERDEN [24,§20]), *i.e. the relation* "dep" *has the following three properties:*

(Dep 1) G_i dep $\{G_1,\ldots,G_k\}$,

(Dep 2) G dep $\{G_1,\ldots,G_k\}$, G_i dep $\{H_1,\ldots,H_r\}$ *imply* G dep $\{H_1,\ldots,H_r\}$,

(Dep 3) Exchange condition. G dep $\{G_1,\ldots,G_k\}$, *not* G dep $\{G_1,\ldots,G_{k-1}\}$ *imply* G_k dep $\{G_1,\ldots,G_{k-1},G\}$.

PROOF. The first two conditions are straightforward. For the proof of (Dep 3) we first remark $[G_1,\ldots,G_{k-1}] < [G_1,\ldots,G_k]$, hence $\mathrm{Dim}[G_1,\ldots,G_k] = \mathrm{Dim}[G_1,\ldots,G_{k-1}] + 1$. For the same reason $\mathrm{Dim}[G_1,\ldots,G_{k-1},G] = \mathrm{Dim}[G_1,\ldots,G_{k-1}] + 1$. Due to $[G_1,\ldots,G_{k-1},G] \le [G_1,\ldots,G_{k-1},G_k]$ and (1.9) we have equality for these spaces, hence $G_k \le [G_1,\ldots,G_{k-1},G]$. □

DEFINITION 3.2. As usual we say $\{G_1,\ldots,G_k\} \subseteq G_x$ is *independent* iff G_i does not depend on $\{G_1,\ldots,G_{i-1},G_{i+1},\ldots,G_k\}$ for all $i \in \{1,\ldots,k\}$. Otherwise the set is called *dependent*.

THEOREM 3.2. *Let F be a finite nearaffine space and* $G_1,\ldots,G_k$ *straight lines of F going through a fixed point. Then*

$$\mathrm{Dim}[G_1,\ldots,G_k] \le k \tag{3.3}$$

and equality holds iff $\{G_1,\ldots,G_k\}$ *is independent.*

PROOF. Formula (3.3) follows easily from (3.1), (1.5) and definition 1.2 by induction on k. Equality in (3.3) is equivalent with

$$G_{i+1} \not\le [G_1,\ldots,G_i] \tag{3.4}$$

for all $i \in \{1,\ldots,k-1\}$, hence by proposition 3.1, equivalent with

$$G_{s_{i+1}} \not\le [G_{s_1},\ldots,G_{s_i}] \tag{3.4'}$$

for any permutation $\{s_1,\ldots,s_k\}$ of $\{1,\ldots,k\}$. Now it is easy to see that the specialization i = k is really equivalent to (3.4'). This completes the proof. □

COROLLARY. *In a finite nearaffine space of dimension* d *the set* G_x *of all straight lines through* x *contains* d *independent straight lines, but more than* d *straight lines of* G_x *are always dependent.*

THEOREM 3.3. *Any* k*-dimensional* ($k \geq 1$) *subspace* U *of a finite nearaffine space containing the point* x *can be represented by*

$$U = [G_1, \ldots, G_k],$$

where $G_1, \ldots, G_k$ *are independent straight lines of* G_x.

PROOF. We prove this theorem by induction on k = Dim U. For k = 1 we have U = [U]. For k > 1 let V be a maximal proper subspace of U. It has the dimension k-1. Let $G_k \leq U$, $G_k \not\leq V$ a straight line, then $U = [V, G_k]$ and the theorem is proved by using the induction hypothesis. □

PROPOSITION 3.2. *Let* $U = [G_1, \ldots, G_k]$, $G_i \in G_x$ *and* $y \in U$. *Define* $G_i' := (y \parallel G_i) \in G_y$. *Then also* $U = [G_1', \ldots, G_k']$.

PROOF. Due to $y \in U$ and $G_i \leq U$ we have $G_i' \leq U$ by proposition 1.2. This implies $[G_1', \ldots, G_k'] \leq U$. But otherwise $G_i = (x \parallel G_i')$ and the same argument yields the converse inequality. □

PROPOSITION 3.3. *Let* U, G_i, y *and* G_i' *be defined as in proposition 3.2 and let*

$$U_y := [G_1', \ldots, G_k'].$$

Then either $U_y = U$ *or* $U \cap U_y = \emptyset$.

PROOF. Assume $z \in U \cap U_y$. Then proposition 3.2 yields $U = U_z = U_y$. □

4. PENCILS OF PARALLEL HYPERPLANES

Let $F = (X, L, \sqcup, \parallel)$ be a finite nearaffine space of order n and dimension d. Let H be an arbitrary hyperplane (cf. definition 1.3) of F; it has the dimension d-1. By (G2) (cf. chapter I,section 3) there is a straight line G such that $H \cap G = \{x\}$ is a point. Using theorem 3.3 and proposition

3.2 it is easy to see that the hyperplanes H_y (for the definition cf. proposition 3.3) are pairwise disjoint for different $y \in G$. By counting arguments we thus obtain

$$(4.1) \qquad X = \dot{\bigcup_{y \in G}} H_y$$

where the dot on the union sign means that the sets whose union will be formed are pairwise disjoint. The set

$$(4.2) \qquad \mathcal{H} := \{H_y \mid y \in G\}$$

is called the *pencil of* (parallel) *hyperplanes generated by* H *) or more briefly a *pencil*. We shall give some properties of such a pencil. For most of them we essentially use the finiteness of the space F. The first proposition, however, is also valid in the infinite case provided of course that hyperplanes do exist.

LEMMA 4.1. *Let* H *be a hyperplane,* $x \notin H$ *and* L *a line with* x *as a base point and* $y \in L\backslash\{x\}$. *Moreover, let* $G \not\parallel H$ *be straight and* $\bar{x},\bar{y}$ *the projections of* x,y *resp. in the direction* G (cf. (1.4)). *Assume* $\bar{x} = \bar{y}$ *or* $L \nparallel \bar{x}\bar{y}$. *Then* $L \cap H \neq \emptyset$.

PROOF. This is trivial if $L = G$. Assume therefore $L \neq G = x\bar{x}$. Then $\bar{x} \neq \bar{y}$. Using (Pa) (cf. chapter I,theorem 3.1), we obtain

$$(y \parallel \bar{y}\bar{x}) \cap (x\bar{x}) \neq \emptyset;$$

hence by (G1)

$$(y \parallel \bar{y}\bar{x}) \cap (x\bar{x}) = \{x'\}.$$

Now $L \neq G$ implies $x' \neq y$ and $xy \nparallel \bar{x}\bar{y}$, hence $yx \nparallel \bar{y}\bar{x}$ by (P3); this implies $x' \neq x$. Thus we can apply the Veblen condition (V) and we get

$$(\bar{x}\bar{y}) \cap (xy) \neq \emptyset.$$

---

*) It is easy to verify that the pencil is independent of the special position of G provided only that $G \not\parallel H$.

This proves the lemma because of $\bar{x} \sqcup \bar{y} \subseteq H$. □

PROPOSITION 4.1. *Let* H *and* H' *be hyperplanes of the same pencil and* L *be a line contained in* H. *Then* $y \in H'$ *implies* $(y \parallel L) \subseteq H'$.

PROOF. This is clear for $H = H'$ by proposition 1.2. Assume $H \neq H'$. Let x be a base point of L and let $x' \in H'$ be such that $x \sqcup x'$ is straight. Assume $z \in L \setminus \{x\}$ and let $\bar{z}$ be the projection of z onto H' in the direction G (cf. (1.4)). Obviously $x \sqcup \bar{z} \subseteq H'$. If $x' \sqcup \bar{z} \pitchfork x \sqcup z = L$ then $(x' \sqcup \bar{z}) \cap H \neq \emptyset$ according to lemma 4.1, hence $H \cap H' \neq \emptyset$. But this contradicts $H \neq H'$ and (4.1). Thus $H' \supseteq (x' \parallel L) = x' \sqcup z \parallel L$ and also $(y \parallel L) \subseteq H'$. □

PROPOSITION 4.2. *Let* L *be a line and* H *a pencil of hyperplanes. Then either* $L \subseteq H$ *for exactly one* $H \in \mathsf{H}$ *or* $|L \cap H| = 1$ *for all* $H \in \mathsf{H}$.

PROOF. Assume $L \not\subseteq H$ for all $H \in \mathsf{H}$. Let x be a base point of L and H_0 the (by (4.1)) uniquely determined hyperplane of H containing x. By our hypothesis there exists an $H_1 \in \mathsf{H} \setminus \{H_0\}$ and a point $y \in L \cap H_1$, consequently $L = x \sqcup y$. Let G be a straight line containing x but not contained in H_0. Using (4.1) we see that $\{x'\} := G \cap H_1$ defines a point. Let H_2 be an other hyperplane of H and x" defined by $\{x''\} := G \cap H_2$. If $y = x'$ then $L = G$ has exactly one point in common with H_2. If $y \neq x'$ then $x' \sqcup y \subseteq H_1$. According to proposition 4.1 we get $(x'' \parallel x' \sqcup y) \subseteq H_2$. Now (V) implies $(x \sqcup y) \cap (x'' \parallel x' \sqcup y) \neq \emptyset$. Hence $L \cap H_2 \neq \emptyset$ for all $H_2 \in \mathsf{H}$. But both H and L contain n elements. By counting arguments thus $|L \cap H_2| = 1$ for all $H_2 \in \mathsf{H}$. This proves the proposition. □

COROLLARY. *Any hyperplane* H *of a finite nearaffine space* $F = (X, L, \sqcup, \parallel)$ *is a flat* (cf. ANDRÉ [1]), *i.e. any line incident with* $x, y \in H$, $x \neq y$ *(not necessarily as base points) lies completely in* H.

We now generalize this property.

THEOREM 4.1. *Every subspace of a finite nearaffine space is a flat.*

PROOF. Let U be a subspace. We use induction on $s_U := |X| - |U|$. For $s_U = 0$ we have $X = U$ and the theorem is true. Assume $s_U > 0$ and select V in such a way that U is a maximal subspace of V, i.e. a hyperplane of the space F_V (cf. (1.2)); this can be done e.g. by using proposition 1.3. Obviously $s_V < s_U$ so that we can apply the induction hypothesis to V. Let x,y be dif-

ferent points of U and L a line incident with both x and y. The induction hypothesis and $x,y \in V$ imply $L \subseteq V$, hence L is a line belonging to the space F_V. The corollary to proposition 4.2 now proves the theorem. □

We apply this theorem and obtain a generalization of proposition 3.3 of chapter I.

PROPOSITION 4.3. *Let* U *be a subspace,* $x \in U$ *and* L *a line having* x *as a base point such that* $L \not\subseteq U$. *Then* $y \in U\backslash\{x\}$ *implies* $L \cap (y \vee L) = \emptyset$.

PROOF. Assume $z \in L \cap (y \vee L)$. Then $z \notin U$ and $L = x \sqcup z$ and $(y \vee L) = y \sqcup z$. Due to (P3) we get

$$z \sqcup x \vee L \vee (y \vee L) \vee z \sqcup y ,$$

hence $z \sqcup x = z \sqcup y$ by (P1). Because of $x,y \in U$, $x \neq y$, theorem 4.1 now yields $z \sqcup x \subseteq U$, hence $z \in U$ contradicting our hypothesis. □

Now we give another description of the hyperplanes of a pencil.

PROPOSITION 4.4. *Let* H *be a hyperplane and* y *a point. Then that hyperplane of the pencil defined by* H *and passing through* y *can be described by*

$$\bigcup_{H \supseteq L \in \mathcal{L}} (y \vee L) .$$

PROOF. Let H' be the (uniquely determined) hyperplane going through y and belonging to the pencil generated by H. Due to proposition 4.1 we have

$$\bigcup_{H \supseteq L \in \mathcal{L}} (y \vee L) \subseteq H' .$$

But on the other hand both point sets have the same number of points, namely n^{d-1}. Hence they coincide. □

We conclude this section with a characterization of two-dimensional subspaces.

DEFINITION 4.1. A nearaffine space is called a *nearaffine plane* if every line L intersects every straight line $G \nparallel L$ in exactly one point (cf. also ANDRÉ [2]).

PROPOSITION 4.5. *A nearaffine space is a nearaffine plane iff its dimension is two.*

PROOF.

(1) Dim X = 2. Then any straight line G is a hyperplane. Using proposition 4.2 we see that any line L $\not\parallel$ G has exactly one intersection point with G.

(2) L $\nparallel$ G *implies* $|L \cap G| = 1$. Assume Dim X $\geq$ 3. The corollary of theorem 3.2 then implies the existence of three independent straight lines G_1, G_2, G_3 going through a point x. According to the dependence conditions stated in section 3 we assume $G = G_1$ without restriction of generality. But then e.g. $(y \parallel G_2)$ has no intersection with G whenever $y \in G_3 \setminus \{x\}$. □

5. WEAK SUBSPACES

Let $F = (X, L, \sqcup, \parallel)$ be a finite nearaffine space of order n.

DEFINITION 5.1. A subset $S \subseteq X$ such that $x,y \in S$, $x \neq y$, imply $x \sqcup y \subseteq S$ is called a *weak subspace*.

Trivially any subspace is a weak subspace. It is our aim to prove the converse relation, i.e. that both concepts coincide.

By proposition 1.2 any subspace is also a weak subspace. It is our aim to prove the converse relation, i.e. that both concepts coincide.

PROPOSITION 5.1. *The intersection of weak subspaces is also a weak subspace.*

For the sake of simplicity we shall give the two following definitions.

DEFINITION 5.2. Let U be a subspace and L a line. We say L is *parallel* to U if for any $x \in U$ we have $(x \parallel L) \subseteq U$. The relation thus defined is denoted by

$$L \parallel U. \tag{5.1}$$

Moreover, a weak subspace satisfying (S2) (cf. definition 1.1), i.e. a subspace, is sometimes also called a *strong subspace*.

THEOREM 5.1. *Let* $F = (X,L,\sqcup,||)$ *be a finite nearaffine space of order* $n \geq 3$ *and let* U *be a set of points. Then the following conditions are equivalent:*

(A) $x,y \in U$, $x \neq y$ *imply* $x \sqcup y \subseteq U$ *(i.e.* U *is a weak subspace).*

(B) U *is a subspace* (cf. def. 1.1).

(C) U *is a flat, i.e. any line incident with* $x,y \in U$, $x \neq y$ *completely lies in* U.

PROOF. (C) $\Rightarrow$ (A) is trivial and (B) $\Rightarrow$ (C) is theorem 4.1. It remains to prove (A) $\Rightarrow$ (B). We subdivide this proof into several steps*).

(1) Minimal counterexample. Assume there is a finite nearaffine space possessing a proper weak subspace (i.e. a set of points being a weak subspace but not a strong one). Then there exists a nearaffine space say $F = (X,L,\sqcup,||)$ of minimal dimension d with this property. In a plane the only weak subspaces are the empty set, the points and the straight lines, all of them being strong. Hence $d \geq 3$.

(2) Minimal weak subspaces. Let S be a fixed proper weak subspace of F with minimal $|S|$. Moreover, let U be a fixed proper subspace of S with maximal $|U|$. By the minimality of S and $U \subset S$ we see U is strong.

(3) *If* G *is straight,* $G \subset S$, *then* $G \,||\, U$; *especially* U *is not a point.* Assume $G \not|| U$ and let be $x \in U$ and $G' := (x \,||\, G)$. Then $G' \cap U = \{x\}$. We have to prove $G' \subseteq S$. If $x \in G$ we have $G' = G \subseteq S$; assume therefore $x \notin G$. Since $n \geq 3$ we can chose three pairwise different points $r,s,t \in G$. Now (Pa) implies the existence of a $p \in (t \,||\, s \sqcup x) \cap (x \,||\, G)$. By proposition 3.3 of chapter I we have $p \neq x$. Now (V) implies that $q \in (r \sqcup x) \cap (t \,||\, s \sqcup x)$ exists. Moreover, $q \in S$ by $r \sqcup x \subseteq S$. If $q = t$ then $x \in G$, hence $q \neq t$ and $(t \,||\, s \sqcup x) = t \sqcup q \subseteq S$, whence $p \in S$ and thus $G' = x \sqcup p \subseteq S$. Because U is strong also $[U,G']$ is strong by proposition 1.3. On the other hand $U \subset [U,G'] \subseteq S$, hence $[U,G'] = S$ due to the maximality of $|U|$ assumed in (2). But then S is strong, contradicting our hypothesis on S.

(4) *Let* V *be strong,* $V \supseteq U$. *Then either* $V \cap S = U$ *or* $V = X$. If $V \cap S \neq U$ then $U \subset V \cap S \subseteq S$, hence $V \cap S$ is not strong by (2). This implies $S \subseteq V$ and by (1) it must be $V = X$.

---

*) I could only prove this theorem with the further hypothesis $L \subseteq U$, $x \in U \Rightarrow (x \,||\, L) \subseteq U$ (cf. the 1st ed.). I owe to O. BACHMANN this proof working with weaker hypotheses.

(5) *Let* V *be strong,* $V \supseteq U$ *and* $V \cap S \subseteq U$ *such that* V *has maximal dimension. Then* V *is a hyperplane.* For $V \cap S \subseteq U$ first implies $V < X$. Let $x \in U$. Then we can find a straight line G with $G \cap V = \{x\}$, especially $G \pitchfork U$. Hence $G \not\subseteq S$ by (3). Due to $U \leq V < [V,G]$ and the maximality of Dim V we have $[V,G] \cap S \not\subseteq U$, hence $[V,G] = X$ by (4). Thus V is a hyperplane.

(6) *Let* G *be straight and* $x \in U \cap G$, $G \cap V = \{x\}$. *Then there exists a hyperplane* W *such that* $G \subseteq W$ *and* $U \not\subseteq W$. Step (1) implies $d = \text{Dim } X \geq 3$, hence $\text{Dim } V \geq 2$ by (5). Let G' be straight such that $x \in G' \subseteq U$. Using the results of section 3, esp. theorems 3.1 and 3.3, there exist independent straight lines $G' = G_1,\ldots,G_{d-1}$ such that $V = [G_1,\ldots,G_{d-1}]$. Then $W := [G_2,\ldots,G_{d-1},G]$ is a hyperplane with the desired properties.

(7) *Let* x,U,G,V,W *and* S *be as before. If* L *is a line such that* $L \subseteq S$ *but* $L \not\subseteq V$ *and such that a base point* b *of* L *is in* U, *then* $L \parallel W$. Assume $L \not\parallel W$ and $y \in U \backslash W$, hence also $S \not\subseteq W$. The propositions 4.1 and 4.2 imply

$$(y \parallel L) \cap W = \{z\}$$

for one point z. We have to show $z \in S$. For this select a point $y' \in U$ such that $y' \sqcup b$ is straight; this is possible because U is not a point by (3). Chose $h \in (y' \sqcup b) \backslash \{y',b\}$ and $b' \in L \backslash \{b\}$. Then (V) implies the existence of a point $d \in (y' \parallel L) \cap (h \sqcup b')$ with $d \neq y'$. From $h,b' \in S$ it follows $d \in S$, hence $(y' \parallel L) = y' \sqcup d \subseteq S$. Applying (G2) on U we conclude $(y \parallel L) \subseteq S$ for all $y \in U$, hence $z \in S$ and then $z \in S \backslash V$ due to $L \not\subseteq V$. Because of $S \cap W \subset S$ we know that $S \cap W$ is strong, hence there exists a chain of straight lines totally contained in $S \cap W$ and connecting x with z. But this contradicts (3). The hypothesis $L \not\parallel W$ is thus wrong.

(8) Conclusion of the proof. Let again be $y \in U \backslash W$. As before let be $x \in U \cap W$ and let L be a line in S having x as a base point such that $L \not\subseteq U$. By (7) with b = x we have $L \parallel W$, hence $L \subseteq W$ by (P1). Now let be $z \in L \backslash \{x\} \subseteq W$ and $L' := y \sqcup z \subseteq S$, hence $L' \parallel W$ by (7). Using (P3) we obtain

$$z \sqcup y = (z \parallel L') \subseteq W,$$

again by (7). But this gives $y \in W$, contradicting the hypothesis $y \notin W$. □

COROLLARY. *The set* U *of all subspaces forms a lattice with respect to the set-theoretical intersection* $\cap$ *and the generating union given by*

$$[U,V] = \bigcap_{U,V \subseteq W \in \mathcal{U}} W, \qquad (U,V \in \mathcal{U}).$$

The same is also true for the set $\mathcal{U}_x$ *of all subspaces containing the point* x.

PROPOSITION 5.2. *Let* U *be a non-void subspace. Then*

$$U = \bigcap\{H \mid U \subseteq H, H \text{ hyperplane}\}. \tag{5.2}$$

PROOF. Let D be the intersection of all hyperplanes $H \supseteq U$. Then trivially $D \supseteq U$. Assume now $x \in U$ and $y \in D \setminus U$. Then $V := [U,y] \supset U = [G_1,\ldots,G_s]$ where the straight lines $G_i \in \mathcal{G}_x$ are assumed to be independent. There exists a straight line $G_{s+1} \leq V$ not lying in U. Enlarge the set $\{G_1,\ldots,G_{s+1}\}$ to a set $\{G_1,\ldots,G_d\}$ of d independent straight lines of $\mathcal{G}_x$ ($d = \mathrm{Dim}\, F$). Then $H := [G_1,\ldots,G_s,G_{s+2},\ldots,G_d]$ is a hyperplane with $U \leq H$ but $V \not\leq H$, hence $D \not\leq H$, contradicting the construction of D. □

6. NEARAFFINE SPACES AS COMBINATORIAL DESIGNS

Let $F = (X,\mathcal{L},\sqcup,\|)$ be a finite nearaffine space of order n. We shall prove in this section that given any two different points x and x' such that $x \sqcup x'$ is not straight there exist exactly n different lines incident with both x and x' (cf. also ANDRÉ [2], BACHMANN [8]).

PROPOSITION 6.1. *Let* x *and* x' *be different points. Then there exists a pencil* H *of parallel hyperplanes* (cf. section 4) *such that* $x \in H \in \mathcal{H}$, $x' \in H' \in \mathcal{H}$ *and* $H \neq H'$.

PROOF. According to proposition 5.2 applied to $U = \{x\}$ there exists a hyperplane H containing x but not x'. The pencil generated by H (cf. section 4) has the desired property. □

PROPOSITION 6.2. *Let* U *be a subspace and* x,x' *two different points in it. Let* $L,L' \not\leq U$ *be lines having* x,x' *resp. as base points. Then* $|L \cap L'| \leq 1$.

PROOF. First $L,L' \not\in U$ imply $L \cap U = \{x\}$ and $L' \cap U = \{x'\}$. Assume $L \cap L'$ contains two different points y and y'. Both these points do not belong to U. Using (P3) we obtain

$$y \sqcup x' \;||\; x' \sqcup y = x' \sqcup y' \;||\; y' \sqcup x' .$$

By counting arguments we conclude that the set S of all those points $z \in (x \sqcup x') \setminus \{x\}$ for which there exists a line L^* with a base point on L and parallel to $y \sqcup x'$ such that $L^* \cap (x \sqcup x') = \{z\}$ is a proper subset of $x \sqcup x'$. For $z^* \in (x \sqcup x') \setminus S$ we then have

$$(z^* \;||\; x' \sqcup y) \cap L = \emptyset$$

contradicting the Veblen condition (V). □

PROPOSITION 6.3. *Let* x *and* x' *be two different points of a finite nearaffine space of order* n. *Then there exist at most* n *lines incident with both* x *and* x'.

PROOF. Let H be a pencil as in proposition 6.1 and H, H' those (uniquely determined) hyperplanes of H going through x, x' resp. For any $H^* \in H \setminus \{H,H'\}$ there exists at most one point y such that $y \sqcup x = y \sqcup x'$; this is a consequence of the preceding proposition. But H possesses exactly n hyperplanes (cf. (4.1),(4.2)). This proves the proposition. □

THEOREM 6.1. *Two different points on a finite nearaffine space of order* n *not being on a straight line are incident with exactly* n *lines.*

PROOF. Let x be a fixed point of the nearaffine space F considered as incidence structure (with respect to ϵ). Denote by F_x the internal structure (cf. DEMBOWSKI [10,p.3]) obtained from F by removing x and all lines not incident with x. Let v', k' and b' be the number of all points, all points on a line and all lines resp. of F_x. Then theorem 2.1 gives

$$(6.1) \qquad \begin{cases} v' = v-1 = n^d-1, \\ k' = k-1 = n-1, \\ b' = r = \gamma + n\left(\dfrac{n^d-1}{n-1} - \gamma\right) = n\dfrac{n^d-1}{n-1} - (n-1)\gamma; \end{cases}$$

here γ is the number of all straight lines of F incident with a given point

(cf. proposition 2.3 and the remark thereafter). Let v_i' ($i \in \mathbb{N}$) be the number of all points of F_x incident with exactly i lines of F_x. Clearly

$$v_1' = \gamma(n-1) \tag{6.2}$$

is the number of all those points y of F_x for which $x \sqcup y$ is straight. By proposition 6.3 we have $v_i' = 0$ for $i > n$. Trivially

$$n^d - 1 = v' = \sum_{i=1}^{n} v_i'. \tag{6.3}$$

Using the principle of *double counting* (see e.g. DEMBOWSKI [10,p.4,(10)]) we get

$$b'k' = \sum_{i=1}^{n} iv_i'. \tag{6.4}$$

Assume now $v_i' > 0$ for some $i \in \{2,\ldots,n-1\}$. Using (6.1) to (6.4) this hypothesis would give the contradiction

$$b'k' < v_1' + n \sum_{i=2}^{n} v_i' = n \sum_{i=1}^{n} v_i' - (n-1)v_1' =$$

$$= n(n^d-1) - \gamma(n-1)^2 = b'k'.$$

This proves the theorem. □

DEFINITION 6.1. Let x,x' be two different points of a finite nearaffine space $F = (X,L,\sqcup,||)$. Define $x \nabla x'$ as the set of all points being base points of lines through x and x', i.e.

$$x \nabla x' := \{x,x'\} \cup \{y \in X \mid y \sqcup x = y \sqcup x'\}. \tag{6.5}$$

Point sets of this type are called *gravity curves* *) *(Schwerpunktkurven)* or *blocks* (cf. ANDRÉ [2,II,§3]). The points x,x' are called the *nodes (Knoten)* of $x \nabla x'$.

*) This name has been chosen because in the case of a nearfield space (cf. chapter I,definition 3.3) the set defined by (6.5) takes the form

$$x \nabla x' = \{y \mid \alpha y + \alpha' y = \alpha x + \alpha' x' \text{ for some } (\alpha,\alpha') \in F^2 \setminus \{(0,0)\}\}. \tag{6.5'}$$

Here $y \in x \nabla x'$ looks like the centre of gravity (Schwerpunkt) of x and x' provided that x,x' are covered by suitable "masses" α,α' resp. (not both 0). Of course this interpretation here is only a formal one and has no physical reality

We note the following obvious results on gravity curves (see also ANDRÉ [2]).

THEOREM 6.2. *Let* $F = (X,L,\sqcup,\|)$ *be a finite nearaffine space of order* n. *Then the join* ∇ *defined by* (6.5) *has the following properties:*

(1) $x,x' \in x\nabla x'$,

(2) $x\nabla x' = x'\nabla x$,

(3) $x\nabla x'' = x\nabla x',\ x'' \neq x' \Rightarrow x\nabla x' = x'\nabla x''$,

(4) $|x\nabla x'| = n$,

(5) $x\nabla x' = x\sqcup x'$ *iff* $x\sqcup x'$ *is straight*,

(6) *if* H *is a hyperplane not containing* $x\sqcup x'$ *then* $|H \cap (x\nabla x')| = 1$,

(7) *if* U *is a subspace,* $x,x' \in U$ *and* $x \neq x'$, *then* $x\nabla x' \subseteq U$.

III. FINITE NEARAFFINE SPACES WITH SPECIAL PROPERTIES

In this chapter we expand the theory developed in the last chapter to special types of finite nearaffine spaces.

1. THE LITTLE DESARGUES CONFIGURATION AND TRANSLATIONS

It is our aim to prove the validity of the little Desargues configuration (D1) (cf. chapter I,section 4) for all finite nearaffine spaces of dimension ≥ 3. To do so it is useful to generalize (D1) in the following way.

(D1') *If* $x,x',y,y',z,z' \in X$ *are pairwise different and* $x\sqcup x' \parallel y\sqcup y' \parallel z\sqcup z'$ *are pairwise different lines then* $x\sqcup y \parallel x'\sqcup y'$ *and* $x\sqcup z \parallel x'\sqcup z'$ *imply* $y\sqcup z \parallel y'\sqcup z'$.

REMARK. (D1') yields (D1) by the additional hypothesis $x\sqcup x'$ is straight.

PROPOSITION 1.1. (D1') *holds provided that there exists a hyperplane* H *containing* x,y,z *but not* $x\sqcup x'$ *(and hence also not* $y\sqcup y'$ *and* $z\sqcup z'$*).*

PROOF. The hyperplane H' through x' belonging to the pencil (cf. chapter II, section 4) generated by H is different from H due to the hypothesis. Using

$x\sqcup y \parallel x'\sqcup y'$, $x\sqcup z \parallel x'\sqcup z'$ and proposition 4.4 from chapter II, we get $y',z' \in H'$. Applying now propositions 4.1 and 4.2 (chapter II) we get $x\sqcup z \parallel x'\sqcup z'$, hence (D1'). □

For the proof of the general validity of (D1) we need the following lemma.

LEMMA 1.1. *Let* G *be a straight line and* $x,y \notin G$. *Then there exists a hyperplane* $H \geq G$ *not containing* x *and* y.

PROOF. Define $U := [G,x,y]$. Applying chapter II, proposition 5.2 to G on the one hand and to U on the other hand we immediately obtain a hyperplane with the desired properties. □

THEOREM 1.1. *In a finite nearaffine space* F *with dimension* $d \geq 3$ *the little Desargues configuration* (D1) *generally holds.*

PROOF. This is true if x,y,z and x',y',z' lie on different hyperplanes by proposition 1.1. Assume now that all points occurring in (D1) belong to the same hyperplane H which due to $d \geq 3$ is at least two-dimensional. Applying lemma 1.1 we can find a hyperplane S of H containing x and x' but not y and z. Let G be a straight line through x with $G \nsubseteq H$. Then $[S,G] =: H_1$ is another hyperplane of our space F. Fix $x'' \in G\backslash\{x\}$ and let H' be that hyperplane of the pencil generated by H which contains x''. Using results of section 4 from chapter II we obtain y'' and z'' by

$$\{y''\} := (y \parallel G) \cap H', \quad \{z''\} := (z \parallel G) \cap H'$$

resp. and $x\sqcup y \parallel x''\sqcup y''$, $x\sqcup z \parallel x''\sqcup z''$ and $y\sqcup z \parallel y''\sqcup z''$. Now we use proposition 1.1 for the triangles x,x',x'' and y,y',y'' and conclude $x''\sqcup x' \parallel y''\sqcup y'$, similarly $x''\sqcup x' \parallel z''\sqcup z'$. Again applying proposition 1.1, now to the triangles x'',y'',z'' and x',y',z', we get the desired result $y\sqcup z \parallel y'\sqcup z'$. □

We apply this theorem in view of group-theoretical properties.

THEOREM 1.2. *The translations* (cf. chapter I, definition 2.1) *of a finite at least three-dimensional nearaffine space form an abelian group generated by the straight translations* (cf. chapter I, definition 2.2) *and acting sharply transitive on the points of the space.*

SKETCH OF PROOF.*) If x,x' are points such that $x \perp x'$ is straight then due to (D1) there exists a straight translation mapping x onto x'. The proof goes just as in the case of the affine space (see e.g. ARTIN [1,chap. II]). According to (G2) we now see that the translations form a group generated by the straight translations. This group operates transitively on the points. On the other hand it is easy to verify that any translation is uniquely determined by its action on one point. Hence the transitivity is sharp.

It remains to prove that the translation-group is abelian. To do so we first show by straightforward methods using (Pa) and (G1) that two straight translations commute if they have different "directions". Using this result one verifies this commutativity also for straight translations with the same direction. Due to the fact that the straight translations generate the whole group the commutativity is now straightforward.

By standard methods we can now shift the structure of the translation-group to the points. In this way X becomes an abelian group whose operation and neutral element are, as usual, denoted by + and 0 resp. The translations are the mappings of the type

$$x \mapsto x+v, \qquad (x,v \in X). \tag{1.1}$$

PROPOSITION 1.2. *A subspace* U *containing* 0 *is a subgroup of* (X,+).

PROOF. Let $u \in U$ and define

$$S-u := \{x-u \mid x \in S\} \tag{1.2}$$

for any subset S of X. By the definition of a translation (chapter I, definition 2.1) we have

$$L \parallel L-u \tag{1.3}$$

for all lines L. Hence $L \subseteq U$ implies $L-u \subseteq U$ due to $u \in U$ and chapter II, proposition 1.2. But

$$U = \bigcup \{L \mid 0 \in L \subseteq U,\ L \text{ line}\}$$

---

*) A detailed proof of this theorem will be given in ANDRÉ [5].

yields now $U-u \subseteq U$, hence U is a subgroup of X. □

LEMMA 1.2. $U \in U_0$ *and* $G \in G_0$ *imply* $U+G = [U,G] \in U_0$.

PROOF. First U, G and [U,G] are subgroups of X. Hence $U+G \subseteq [U,G] = \bigcup_{u \in U} (u \parallel G)$ (cf. chapter II,(1.3)). On the other hand let $v \in [U,G]$ and $\bar{v}$ its projection on U in the direction G (cf. chapter II,(1.4)). Then $x \mapsto x + (v-\bar{v})$ is a translation with direction G and $v = \bar{v} + (v-\bar{v}) \in U+G$. □

By induction we easily see using this lemma

PROPOSITION 1.3. $U,V \in U_0$ *imply* $[U,V] = U+V \in U_0$.

We conclude this section with a description of the lattice U_0 of all subspaces through 0.

THEOREM 1.3. *The lattice* $(U_0,\cap,[])$ (cf. also chapter II, corollary of theorem 5.1) *is complete, of finite length, modular and complementary.* *)

PROOF. The lattice U_0 is complete and of finite length because it is finite. The modularity follows from proposition 1.3 and the fact that the lattice of all subgroups of an abelian group is modular. Due to theorem 3.3 of chapter II any element $\neq \{0\}$ of U_0 can be represented as a join of atoms (i.e. straight lines in our case); this yields that U_0 is complementary.

2. COMPATIBILITY OF STRAIGHT LINES

Let $F = (X,L,\sqcup,\parallel)$ be a finite nearaffine space. We fix a point in X, say 0. Remember that G_0 and U_0 are the sets of all straight lines and all subspaces resp. going through 0.

DEFINITION 2.1. Two straight lines $G,G' \in G_0$ are called *compatible (verträglich)*, symbolically

(2.1) $\quad G' \text{ cp } G,$

*) For these and other lattice-theoretical concepts see e.g. BIRKHOFF [9].

if either G = G' or there exist $x \in G\backslash\{0\}$ and $x' \in G'\backslash\{0\}$ such that $x \sqcup x'$ is straight.

Trivially cp is a reflexive and symmetric relation on G_0. We shall see later that cp is also transitive.

DEFINITION 2.2. A subspace $U \in U_0$ is called *compatible* if any two straight lines $G,G' \leq U$, $\in G_0$ are compatible in the sense of definition 2.1. This obviously means that either U itself is a straight line or the triangle condition (G3) (cf. chapter I,section 3) holds in F_U (cf. chapter II,(1,2)).

PROPOSITION 2.1. *A plane* $E \in U_0$ *is compatible iff any point of* E *contains at least three different lines contained in* E. *)

This is an easy consequence of (P1), (P2) and proposition 4.5 of chapter II.

COROLLARY. G cp G' *implies* [G,G'] *is compatible.*

The following proposition is similar to the diagonal condition (Di) of quadrilaterals (cf. chapter I,section 4).

PROPOSITION 2.2. *Let* x,y,z,w *be pairwise different points not all lying in one plane and such that* $x \sqcup y$, $y \sqcup z$, $z \sqcup w$ *and* $w \sqcup x$ *are straight. Then the diagonals* $x \sqcup z$ *and* $y \sqcup w$ *are straight too.*

PROOF. The intersections $[x \sqcup y, x \sqcup w] \cap [z \sqcup y, z \sqcup w]$ and $[y \sqcup x, y \sqcup z] \cap [w \sqcup x, w \sqcup z]$ are straight lines because x,y,z,w do not belong to the same plane; these straight lines coincide with $x \sqcup z$ and $y \sqcup w$ resp. □

PROPOSITION 2.3. *The diagonal condition* (Di) *holds in a compatible subspace of dimension at least three.*

PROOF. Assume that x,y,z,w are the four corners of the quadrilateral under consideration. If they do not belong to the same plane (Di) holds because of proposition 2.2. Assume that x,y,z,w lie in the same plane $E \leq U$. Due to

---

*) This means that E considered as a nearaffine plane has the type (n,s) with $s \geq 2$ in the sense of ANDRÉ [2,II,§1]; see also section 3 of this chapter.

Dim U ≥ 3 there exists a straight line G ∉ E with x ∈ G ≤ U. Select a u ∈ G\{x} in such a way that u⌊⌋z is straight; this is possible because x⌊⌋z is straight by hypothesis of (Di) and U is compatible. By proposition 2.2 both y⌊⌋u and w⌊⌋u are straight. Further application of proposition 2.2 now to the quadrilateral x,y,u,w yields that y⌊⌋w is straight. □

THEOREM 2.1. *The compatibility relation in the sense of definition 2.1 is transitive. The equivalence classes with respect to this relation are subspaces of* U_0, *hence maximal compatible subspaces, having pairwise only* {0} *as intersection.*

PROOF. Fix a $G \in G_0$ and consider

$$(2.2) \qquad \langle G \rangle := \{G' \in G_0 \mid G' \text{ cp } G\}.$$

If <G> = {G} then all is proved. Let $V \in U_0$ be a subspace with G ≤ V and in which any straight line through 0 and contained in V is compatible to the fixed G. Assume it has already been proved that V is compatible. If V contains all straight lines compatible to G the proof is finished. Otherwise let G' be compatible to G and G' ∉ V. We shall enlarge the subspace V to another compatible subspace. If V = G then [G,G'] is compatible by proposition 2.1. Assume now Dim V ≥ 2 and $0 \in G_1 \leq V$, $G_1 \neq G$. Using proposition 2.2 we easily see G' cp G_1. Hence G' is compatible to all lines of G_0 lying in V. Now select a G" ≤ [V,G'], ∉ V, belonging to G_0 and different from G'. Then [G',G"] is a plane. Using chapter II,(1.3) we see that its intersection with V is a straight line G_2 different from G' and G". Proposition 2.1 yields that G" is compatible to G' and a further use of proposition 2.2 gives the compatibility of G" to all lines of G_0 contained in V. This yields that [V,G'] > V is also compatible.

Repeating this method we finally get a maximal compatible subspace U and the lines of G_0 compatible with the given G are exactly those straight lines which are subspaces of U. Hence cp is an equivalence relation and the lines of G_0 contained in U form the equivalence class generated by G. Of course two different maximal compatible subspaces have only 0 as common element. This proves the theorem. □

THEOREM 2.2. *Let* $\{x_1,\ldots,x_t\} \subseteq U_0$ *be the set of all maximal compatible subspaces of the finite nearaffine space* F = (X,L,⌊⌋,||) *satisfying* (D1). *Then*

the group (X,+) (cf. section 1, esp. theorem 1.2) *is a direct sum of the subgroups* X_i.

PROOF. As subspaces going through 0 the sets X_i are subgroups of X due to proposition 1.2. Moreover,

$$X = \sum_{i=1}^{t} X_i \tag{2.3}$$

because any $G \in G_0$ is contained in (exactly) one X_i, and any $x \in X$ is a sum of elements y such that 0⊔y is straight (i.e. the mapping $z \mapsto z+y$ is a straight translation); this is a consequence of theorem 1.2. It remains to prove that the sum (2.3) is direct, i.e.

$$X_j \cap \sum_{i=1}^{j-1} X_i = \{0\} \tag{2.4}$$

holds for all $j \in \{2,\ldots,t\}$. Assume the contrary of (2.4). Then there exists a j and an x in the intersection stated on the left hand side of (2.4) such that 0⊔x is straight. But theorem 2.1 yields that every straight line lying in X_j not compatible with any straight line on $\sum_{i=1}^{j-1} X_i$ is not contained in this sum. This yields (2.4). □

Let $U_0(X_i)$ be the lattice of all subspaces of X_i ($i\in\{1,\ldots,t\}$) through 0 and $U_0 = U_0(X)$, as before, the lattice of all subspaces through 0.

Theorem 1.3 stated that U_0 forms a complete complementary modular lattice of finite length. By simple lattice-theoretical arguments we are now able to sharpen this theorem.

THEOREM 2.3. *The lattice* U_0 *is the direct product of the lattices* $U_0(X_i)$ *of all subspaces of* X_i *going through* 0 *where the* X_i ($i\in\{1,\ldots,t\}$) *are the maximal compatible subspaces of* X. *The lattices* $U_0(X_i)$ *are irreducible complementary modular lattices of finite length. Especially* U_0 *itself is irreducible iff* X *is compatible, i.e.* (G3) *holds in* X.

We conclude this section with a sufficient condition for the validity of (Di) in a finite nearaffine space.

THEOREM 2.4. *Let* $F = (X,L,\sqcup,||)$ *be a finite nearaffine space in which all*

maximal compatible subspaces X_i *are at least three-dimensional. Then the diagonal condition* (Di) *holds* *) *in* F.

PROOF. This is an obvious consequence of proposition 2.3 and theorem 2.2. □

It is unknown whether the condition on the X_i can be weakened.

3. THE TYPE OF A NEARAFFINE SPACE

Let $\mathcal{F}$ be a finite nearaffine space of order n. As we have seen in chapter II, proposition 2.3 the class of $\mathcal{F}$, i.e. the number γ of straight lines going through a given point in $\mathcal{F}$ does not depend on the special choice of this point. If the space is a plane then define s by $\gamma =: s+1$ and call (n,s) the *type* of this plane (see also ANDRÉ [2,II,§1]). We shall generalize this concept for arbitrary finite nearaffine spaces. The fundamental idea for this is given by the following theorem.

THEOREM 3.1. *All subplanes of a finite compatible nearaffine space (i.e. a space with* (G3); cf. section 2*) have the same type.* **)

PROOF. We may assume that the compatible space under consideration is at least three-dimensional. Let E and E' be two different subplanes; they are also compatible. We have to consider two cases.

(1) $E \cap E' =: G$ is a straight line. Select $x \in G$ and two straight lines $H,H' \neq G$ with $x \in H \leq E$ and $x' \in H' \leq E'$. Fix $y \in H\setminus\{x\}$ and $y' \in H'\setminus\{x\}$ such that $y \sqcup y'$ is straight; this is possible because the space under consideration is compatible. Let K be any straight line through y and contained in E. Define a mapping f by

$$f(K) = \begin{cases} y' \sqcup (K \cap G) \text{ ***)} & \text{if } K \pitchfork G, \\ (y' \parallel K) & \text{if } K \parallel G. \end{cases}$$

---

*) Compatible spaces with (Di) are also called *regular* (cf. ANDRÉ [2,I,§3]).

**) For this theorem no results of section 1 are used.

***) Here we identify the set $K \cap G =: \{z\}$ with the point z.

Due to proposition 2.3 the mapping f is a bijection from the pencil of all straigt lines in E through y onto the respective pencil of the straigt lines in E' through y'. Hence the plane E and E' have the same type.

(2) E,E' are in an arbitrary situation. Then one can find a chain $E = E_0, E_1, \ldots, E_t = E'$ of planes such that $E_i \cap E_{i+1}$ $(i \in \{0,\ldots,t-1\})$ are straight lines and hence E and E' are of the same type. In order to find such a chain we consider the following possibilities.

(a) $E \cap E' = \{x\}$. Select different straight lines G, G' with $x \in G \leq E$ and $x' \in G' \leq E'$. Then $E = E_0$, $E_1 = [G,G']$, $E_2 = E'$ is a chain with the desired properties.

(b) *There are* $x \in E$, $x' \in E'$ *such that* $x \sqcup x'$ *is straight*. Then choose a straight line H such that $(x \sqcup x') \cap H$ is a point. We see $E'' := [x \sqcup x', H]$ is a plane related to E and E' by (a) or (1).

(c) *Arbitrary case*. Select $x \in E$ and $x' \in E'$. By (G2) these points can be connected by a chain of straight lines $x_i \sqcup x_{i+1}$. Choose a plane $E_i \ni x_i$. Using induction on the length of the chain of straight lines we can reduce this case to (a). □

DEFINITION 3.1. A compatible finite d-dimensional ($d \geq 2$) nearaffine space is said to be of *type* (n,s,d) if all subplanes are of type (n,s). A straight line with n points is said to be of *type* (n,1,1).

PROPOSITION 3.1. *In a nearaffine space of type* (n,s,d) *with* $d \geq 2$ *the number of all straight lines through a fixed point is*

$$\gamma = 1 + s + \ldots + s^{d-1} = \frac{s^d - 1}{s - 1} . \tag{3.1}$$

PROOF. We prove this proposition by induction on d. It is true for $d = 2$ due to the definition of s. Let us consider the general case and assume $d > 2$. Select a hyperplane H, a point $x \in H$ and a straight line $G \not\leq H$ through x. Obviously H is compatible. Fix a $y \in G \setminus \{x\}$. By our induction hypothesis the number of all straight lines in H through x is

$$\gamma' := 1 + s + \ldots + s^{d-2}.$$

Given any straight line G' with $x \in G' \leq H$, then [G,G'] is a plane containing y and $x \in G'$, $G'' \leq H$, $G' \neq G''$ imply $[G,G'] \cap [G,G''] = G$. There exist s straight lines $\neq G$ through y on any such [G,G']. Hence the number of all

straight lines through y is

$$\gamma = 1 + s\gamma' = 1 + s \ldots + s^{d-1}. \quad \square$$

Using well-known structure theorems on irreducible complementary modular lattices (see e.g. BIRKHOFF [9]) we get the following properties of compatible spaces.

THEOREM 3.2. *If the compatible nearaffine space* $F = (X, \mathcal{L}, \sqcup, ||)$ *of type* (n,s,d) *is three-dimensional the irreducible complementary modular lattice* U_0 *is a projective plane of order* s. *If* F *is at least four-dimensional then* U_0 *forms an at least three-dimensional, hence desarguesian projective space. In this case* s *must be a power of a prime.*

It is an open question whether s is a divisor of n in the general case.

Select now a point 0 of the finite d-dimensional nearaffine space $F = (X, \mathcal{L}, \sqcup, ||)$ and consider the maximal compatible subspaces $X_1, \ldots, X_t$ through 0. Assume that X_i is of type (n, s_i, d_i). Then we say that F is of *type*

$$(3.2) \qquad (n, \{(s_1,d_1), \ldots, (s_t,d_t)\}).$$

(The curved brackets { } notify of course that the order of the (s_i,d_i) is inessential.) Using theorem 2.2 we can easily prove the following theorem.

THEOREM 3.3. *A finite* d*-dimensional nearaffine space of a type given by* (3.2) *has the class* (cf. the remark after proposition 2.2 from chapter II)

$$(3.3) \qquad \gamma = \sum_{i=1}^{t} \frac{s_i^{d_i} - 1}{s_i - 1}$$

and its dimension is

$$(3.4) \qquad d = \sum_{i=1}^{t} d_i \, .$$

4. THE STRUCTURE OF DESARGUESIAN NEARAFFINE SPACES

In this section we add to the conditions (D1) and (Di) the Desargues configuration (D2) (cf. chapter I, section 4). But while we have reasonable sufficient conditions for the validity of (D1) and (Di) (cf. sections 1 and 2 resp.) no such condition for (D2) is known to the author. But if we assume (D1), (D2) and (Di) and another condition stated below for the nearaffine space F then a complete survey can be given even in the infinite case; this will establish an analogue of Hilbert's characterization of desarguesian affine spaces. Let us now state the additional condition which is a weakened form of (G3):

(G3') Weakened triangle condition. *For any straight line* G *and* $x \in G$ *there exists a straight line* $G' \neq G$ *with* $x \in G'$ *such that* G *and* G' *are compatible* (cf. section 2), *i.e. there exist* $y \in G\backslash\{x\}$ *and* $y' \in G'\backslash\{x\}$ *such that* $y \sqcup y'$ *is straight.*

REMARK. This condition is e.g. true if the hypotheses of theorem 2.4 hold.

THEOREM 4.1. (Structure theorem for desarguesian nearaffine spaces). *Let* $F = (X, \sqcup, \sqcup, \|)$ *be a nearaffine space with the additional conditions* (D1), (D2), (Di) *and* (G3'). *Then* F *is isomorphic to a nearfield space in the sense of definition 3.3 of chapter I.*

SKETCH OF PROOF.*) Using (D1) and theorem 1.2 we get a group-theoretical structure on X in a natural way. For the sake of brevity we call the elements of $(X,+)$ *vectors*. Define the set Q by

$$Q := \{x \in X \mid 0 \sqcup x \text{ is straight}\} \cup \{0\}. \tag{4.1}$$

Introduce now *scalars* as mappings α from Q into itself such that

$$\alpha 0 = 0 \tag{4.2}$$

and

$$u, v \in Q \text{ implies } \alpha u = \alpha v \text{ or } u \sqcup v \parallel \alpha u \sqcup \alpha v\,. \tag{4.2'}$$

---

*) A detailed proof of this theorem will be given in ANDRÉ [5]. The proof goes essentially the same way as ARTIN's proof (cf.[7]) of the structure theorem for desarguesian affine spaces.

We denote the set of all scalars by F. It is rather difficult to prove that for all $\alpha \in F$ either $\alpha = 0$ or α is a bijection on Q and

(4.3) $\qquad \alpha(u+v) = \alpha u + \alpha v$

provided that $u,v,u+v \in Q$. For this proof (Di) must be used. If $\alpha \in F\backslash\{0\}$ then also $\alpha^{-1} \in F$. Moreover, the product of scalars is again a scalar. Using (D2) one can prove that there is an $\alpha \in F$ mapping $u \neq 0$ onto v iff $v \in 0 \sqcup u$. In this case α is uniquely determined by u and v. A $u \in Q$ is called *dependent* of $U \subseteq Q$ if u can be represented as a linear combination of finitely many elements of U whose coefficients belong to F. For this relation all usual dependence conditions hold (see e.g. VAN DER WAERDEN [24,§20], also chapter II, theorem 3.1) so that Q is a dependence space. Let B be a basis of Q, i.e. an independent set generating Q; such a basis exists because of Zorn's lemma. Let $\langle B \rangle$ be the set of all those $x \in X$ which can be represented as a linear combination of B with coefficients in F. Using (G2) one can prove $\langle B \rangle = X$ and extend the multiplication of the elements of F with those of Q to a general multiplication of scalars with vectors. In this way every $\alpha \in F\backslash\{0\}$ becomes an automorphism of (X,+) thus generalizing (4.3). Moreover, the following cancellation law holds:

(4.4) $\qquad \alpha x = \beta x$ implies $x = 0$ or $\alpha = \beta$.

By heavily using (D2) and also (G3'), (Di) (at one step of the proof) one can show that the lines are just the sets of the form $Fx+y$ with $x \in X\backslash\{0\}$, $y \in X$, and two lines $Fx+y$ and $Fx'+y'$ are parallel iff $Fx = Fx'$. Moreover, $u \in Q\backslash\{0\}$ iff for all $\alpha,\beta \in F$ there exists a $\gamma \in F$ such that

(4.5) $\qquad \alpha u + \beta u = \gamma u$.

Using (4.4) we can define an addition $\underset{u}{+}$ on F by $\alpha \underset{u}{+} \beta := \gamma$. It is easy to see that $(F, \underset{u}{+}, \cdot)$ becomes a nearfield and that (X,+) together with F will be a structure defined in chapter I, (3.2) to (3.4). Hence $\bar{F}$ is isomorphic to a nearfield space. □

REMARK. One can further prove that this algebraic structure is uniquely determined by $\bar{F}$ up to isomorphism and that (D3) holds in $\bar{F}$.

If the desarguesian nearaffine space is finite, d-dimensional and compatible then $X = F^d$ for a suitable nearfield F. If the space is of

type (n,s,d) (cf. section 3) it is easy to see that |F| = n and one can prove s = |K| where K is the *nucleus* of F defined by

$$K := \{\alpha \in F \mid (\xi+\eta)\alpha = \xi\alpha + \eta\alpha \quad \text{for all } \xi,\eta \in F\}.$$

Hence in this case s is a divisor of n.

APPENDIX: UNSOLVED PROBLEMS

(1) State a complete set of geometric axioms characterizing the spaces of the type I_Γ (cf. section 2 of chapter I)! This has only been done if Γ is a group as defined in chapter I, (3.5), i.e. essentially a Frobenius group; the essential result in this case is given by the hypotheses of chapter III, theorem 4.1.

(2) Search for further relations between Γ and I_Γ. Especially it might be possible to describe the full automorphism group of I_Γ (which has of course Γ as a subgroup) and the translations T of I_Γ (cf. chapter I, definition 2.1). Under what conditions is T a transitive subgroup of Γ?

(3) This problem stated in the 1st ed. has been solved: If in a group space I_Γ all points are joinable (cf. chapter I, def. 1.2) then Γ is either imprimitive or double transitive (cf. chapter I, Theorem 2.2 of this edition).

(4) Do (V) and (Pa) imply (T) (cf. sections 1 and 3 of chapter I) at least in the finite case? (chapter I, section 1, consequence (f) together with theorem 3.1 imply the converse.)

REMARK. The Veblen-condition (V) is independent of the axioms (L), (R), (P) and (G) of a nearaffine space given in the plane case; this has been proved by J. VAN DE SCHOOT & H. WILBRINK: *Nearaffine planes I*, to appear in Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen.

(5) What further reasonable geometrical configurations may hold in quasi-affine or nearaffine spaces? Are there generalizations of the Reidemeister-, Thomsen-, Klingenberg- and other configurations (cf. also KLINGENBERG [17])? State relations among them! State connections between transitive permutation groups Γ and configurations in I_Γ!

(6) What further properties hold in quasi-affine, nearaffine and especially nearfield spaces if suitable topological properties of $\sqcup$ and $\|$ are

required? Characterize the spaces over the Kalscheuer nearfields (cf. KALSCHEUER [16]).

(7) Let F be a Kalscheuer nearfield and consider

$$X := \{(\xi_i)_{i\in\mathbb{N}} \mid \xi_i \in F, \sum_{i=1}^{\infty} |\xi_i|^2 < \infty\}$$

(for the absolute value in a Kalscheuer nearfield see KALSCHEUER [16]), leading to generalizations of Hilbert spaces. Search for properties of such spaces, especially with respect to functional analysis! Are there reasonable relations to quantum mechanics?

(8) Let $S = (X,L,\sqcup,||)$ be a quasi-affine (or more specially a nearaffine) space and let G be the set of all straight lines in S. Then the incidence structure $\bar{S} := (X,G,\epsilon)$ is a partial plane (cf. DEMBOWSKI [10,p.9]), i.e. any two different points are incident with at most one line of G. Under what conditions for $\bar{S}$ is it possible to reproduce S from $\bar{S}$ in a unique way?

(9) This problem stated in the first ed. has been solved by O. BACHMANN; the answer is affirmative (cf. chapter II, theorem 5.1, including the footnote, in this edition).

(10) Let F be a finite d-dimensional nearaffine space. It is easy to see that two of its points can be connected by a chain of at most d straight lines. Define A_i ($i \in \{1,\ldots,d\}$) as the set of all those 2-subsets $\{x,y\}$ of different points which can be connected by i but no fewer straight lines. Is X together with these A_i an association scheme of class number d (DEMBOWSKI [10,p.281])? If the answer is affirmative then F would be a partial design (l.c.,p.282). Compute the parameters p_{ij}^h and n_i in this case!

Remark. If $d = 2$ then X with A_1, A_2 is an association scheme of class number 2 and thus a strongly regular graph (cf. ANDRÉ [2,II,§2]; for strongly regular graphs see e.g. GOETHALS & SEIDEL [11]).

(11) For any finite nearaffine space $F = (X,L,\sqcup,||)$ there exists a commutative join ∇ defined by (6.5), chapter II. Some of its properties are given in chapter II, theorem 6.2. State further conditions on ∇, if possible in such a way that the original space F can uniquely be reproduced from $(X,L,\nabla,||)$! Give a complete survey of all possible situations of the nodes of $x\nabla y$!

(12) Let $\mathcal{F}$ be a compatible nearaffine space of type (n,s,d) (cf. chapter III, section 3). Under what conditions does s|n hold?

Remark. This is true if $\mathcal{F}$ is desarguesian (cf. chapter III, section 4).

(13) State sufficient conditions for the validity of (D2) (cf. chapter I, section 4)! Does Dim $\mathcal{F} \geq 3$ suffice in the finite case? Is the following *anti-hyperplane-condition* (H) sufficient?

(H) *If* x *and* y *are different points then* [x,y] *is neither the whole space* X *nor a hyperplane in it.**)

(14) Give examples of *proper* finite nearaffine spaces, esp. planes, being no nearfield spaces, hence not dearguesian, or prove that there cannot exist such spaces!

(15) Let $\mathcal{F} = (X, \mathcal{L}, \sqcup, ||)$ be a finite nearaffine space such that in X is defined an addition + such that (X,+) becomes an abelian automorphism-group of $\mathcal{F}$ generated by straight translations with neutral element 0 (cf. chapter III, esp. theorem 1.2 and (1.1) which show that such group exists if Dim $\mathcal{F} \geq 3$). Define kx ($k \in z$, $x \in X$) as usual. Prove or disprove $kx \in 0 \sqcup x$! (Trivially this is true if $0 \sqcup x$ is straight.)

REMARK. If $kx \in 0 \sqcup x$ holds for all $k \in Z$ then all $x \in X \setminus \{0\}$ have the same order in (X,+) and hence this group is elementary abelian: Assume first $x,y \in X \setminus \{0\}$ such that $0 \sqcup x \neq 0 \sqcup y$, and $0 \sqcup x$ is straight. If $kx \neq 0$, $ky = 0$ for a $k \in Z$ then $k(x-y) = kx \in 0 \sqcup x$. If $x-y \notin 0 \sqcup x$ then $0 \sqcup x$ and $0 \sqcup (x-y)$ would have the two different intersection points 0 and k(x-y) contradicting (G1). Hence $x-y \in 0 \sqcup x$, $y \in 0 \sqcup x$ because $0 \sqcup x$ is straight, thus $0 \sqcup x = 0 \sqcup y$ contradicting the hypothesis $0 \sqcup x \neq 0 \sqcup y$. The case $0 \sqcup x = 0 \sqcup y$ straight can be reduced to the previous case in the usual manner. All straight and thus all translations have the same order. (The proof of theorem 1.3 in chapter III of the 1st ed. is so incorrect.)

ADDED IN PROOF

O. BACHMANN (Bern) informed me that II, theorem 5.1 can be proved without condition (2) of weak subspaces (cf. II, def. 5.1) provided that the order n of the space is ≥ 3. Thus problem (9) is solved affirmatively. In this proof the steps (1), (2), (4), (5) and (6) go as in the proof of II,

theorem 5.4, noted in this paper. The proof of the other steps as it differs will be stated subsequently.

(3) We have to show $G' \subseteq S$. If $x \in G$ we have $G' = G \subseteq S$. Assume, therefore, $x \in G$. Since $n \geq 3$ we can choose three pairwise different points $r,s,t \in G$. Now (Pa) implies the existence of a $p \in (t \| s \sqcup x) \cap (x \| G)$. By I, proposition 3.3 we have $p \neq x$. Now (V) implies $q \in (r \sqcup x) \cap (t \| s \sqcup x)$ exists. Now $q \in S$ by $r \sqcup x \leq S$. Moreover, $q = t$ would imply $x \in G$, hence $q \neq t$ and $(t \| s \sqcup x) = t \sqcup q \subseteq S$, whence $p \in S$ and thus $G' = x \sqcup p \subseteq S$.

(7) We additionally assume that the base point b of L is in U and have to prove $L \| W$. For this we have to show $z \in S$. Select $y' \in U$ such that $y' \sqcup b$ is straight; this is possible by I, §3, (G2). Choose $h \in (y' \sqcup b) \setminus \{y', b\}$ and $b' \in L \setminus \{b\}$. Then (V) implies the existence of a $d \in (y' \| L) \cap (h \sqcup b')$ with $d \neq y'$. From $h, b' \in S$ it follows $d \in S$, hence $(y' \| L) = y' \sqcup d \subseteq S$. Applying (G2) on U we conclude $(y \| L) \subseteq S$ for all $y \in U$, hence $z \in S$.

(8) In this step the conclusion $z \sqcup y = (z \| L') \subseteq W$ is a consequence of (P1) holding on the hyperplane W.

REFERENCES

[1] ANDRÉ, J., *On flats and dual-flats in incidence structures*, in: Atti del Convegno di Geometria Combinatoria e sue Applicazioni, Perugia, 1971, pp.7-15.

[2] ANDRÉ, J., *Some new results on incidence structures*, to appear in Proceedings Academia Lincei, Roma.

[3] ANDRÉ, J., *Eine Kennzeichnung der Dilatationsgruppen desarguesscher affiner Räume als Permutationsgruppen*, to appear in Arch. Math. (Basel), 25.

[4] ANDRÉ, J., *Lineare Algebra über Fastkörpern*, Math.Z. 136 (1974), 295-313.

[5] ANDRÉ, J., *Affine Geometrien über Fastkörpern*, to be submitted to Mitt. Math. Sem. Giessen.

[6] ARNOLD, H.J., *Die Geometrie der Ringe im Rahmen allgemeiner affiner Strukturen*, Vandenhoeck & Ruprecht, Göttingen, 1971.

[7] ARTIN, E., *Geometric algebra*, Interscience, New York etc., 1957.

[8] BACHMANN, O., *Ueber eine Klasse verallgemeinerter affiner Räume*, to appear in Monatshefte für Math.

[9] BIRKHOFF, G., *Lattice theory*, third ed., Amer. Math. Soc., Coll. Publ., New York, 1967.

[10] DEMBOWSKI, P., *Finite geometries*, Springer-Verlag, Berlin etc., 1968.

[11] GOETHALS, J.M. & J.J. SEIDEL, *Strongly regular graphs derived from combinatorial designs*, to appear in Canad. Math. J.

[12] HUGHES, D. & F. PIPER, *Projective planes*, Springer-Verlag, Berlin etc., 1973.

[13] JORDAN, P., *Algebraische Betrachtungen zur Theorie des Wirkungsquantums und der Elementarlänge*, Abh. Math. Sem. Univ. Hamburg, 18 (1952) 99-119.

[14] JORDAN, P., *Halbgruppen von Idempotenten und nicht-kommutative Verbände*, J. Reine Angew. Math., 211 (1962) 136-161.

[15] JORDAN, P., J. VON NEUMANN & E. WIGNER, *On an algebraic generalization of the quantum mechanical formalism*, Ann. of Math., 35 (1934) 29-64.

[16] KALSCHEUER, F., *Die Bestimmung aller stetigen Fastkörper*, Abh. Math. Sem. Univ. Hamburg, 13 (1940) 413-435.

[17] KLINGENBERG, W., *Beziehungen zwischen einigen affinen Schliessungssätzen*, Abh. Math. Sem. Univ. Hamburg, 18 (1952) 120-143.

[18] PICKERT, G., *Projektive Ebenen*, Springer-Verlag, Berlin etc., 1955.

[19] SPERNER, E., *Affine Räume mit schwacher Inzidenz und zugehörige algebraische Strukturen*, J. Reine Angew. Math., 204 (1960) 205-215.

[20] TAMASCHKE, O., *Projektive Geometrie I*, BI, Mannheim, 1969.

[21] TAMASCHKE, O., *Projektive Geometrie II*, BI, Mannheim, 1972.

[22] VARADARAJAN, V.S., *Geometry of quantum theory I*, van Nostrand, Princeton, 1968.

[23] VARADARAJAN, V.S., *Geometry of quantum theory II*, van Nostrand, Princeton, 1970.

[24] WAERDEN, B.L. VAN DER, *Algebra I*, 7. Aufl., Springer-Verlag, Berlin etc., 1966.

[25] WIELANDT, H., *Finite permutation groups*, Academic Press, New York etc., 1964.

[26] WILLE, R., *Kongruenzklassengeometrien*, Springer-Verlag, Berlin etc., 1970.

CODING THEORY

MATHEMATICAL CENTRE TRACTS 55, 1974, 115-142

WEIGHT ENUMERATORS OF CODES

N.J.A. SLOANE

Bell Laboratories, Murray Hill, New Jersey 07974, USA

ABSTRACT

A tutorial paper dealing with the weight enumerators of codes, especially of self-dual codes. We prove MACWILLIAMS' theorem on the weight distribution of the dual code, GLEASON's theorem on the weight distribution of a self-dual code, some generalizations of this theorem, and then use GLEASON's theorem to show that very good self-dual codes do not exist.

1. INTRODUCTION

We shall mostly consider codes which are *binary* (have symbols from F_2, the field with two elements) or *ternary* (have symbols from F_3, the field with 3 elements). Let F_q^n denote the vector space of all vectors of length n, i.e., having n components, from F_q.

An [n,k] *code* C over F_q is a subspace of F_q^n of dimension k. The vectors of C are called *codewords*. So a binary code is a set of vectors which is closed under addition. A ternary code is closed under addition and under multiplication by -1.

The (*Hamming*) *weight* of a vector $x = (x_1,\ldots,x_n) \in F_q^n$, denoted by wt(x), is the number of non-zero x_i; and the (*Hamming*) *distance* between vectors $x,y \in F_q^n$ is dist(x,y) = wt(x-y).

If every non-zero codeword in C has weight $\geq d$, the code is said to have *minimum weight* d, and is called an [n,k,d] code; n,k,d are the basic parameters of the code. The codewords contain n symbols, and so the *rate* or *efficiency* of the code is $\frac{k}{n}$. Furthermore the code can correct $[\frac{d-1}{2}]$ errors.

The *dual code* $C^{\perp}$ is the orthogonal subspace to C:

$$C^{\perp} = \{u \mid u \cdot v = \sum_{i=1}^{n} u_i v_i = 0, \text{ for all } v \in C\}.$$

$C^{\perp}$ is an $[n,n-k]$ code.

If $C \subset C^{\perp}$, C is called *self-orthogonal*, while if $C = C^{\perp}$ it is called *self-dual*. (See the examples below.)

Let A_i be the number of codewords in C with weight i. Then the set $\{A_0,\ldots,A_n\}$ is called the *weight distribution* of C. It is more convenient to make a polynomial out of the A_i's. The *weight enumerator* of C is

$$W_C(x,y) = A_0x^n + A_1x^{n-1}y + \ldots + A_ny^n =$$

$$= \sum_{i=0}^{n} A_i x^{n-i}y^i = \sum_{u \in C} x^{n-wt(u)}y^{wt(u)}.$$

This is a homogeneous polynomial of degree n in the indeterminates x and y. We could get rid of x by setting $x = 1$, but the theorems are simpler if W is homogeneous.

The weight enumerator gives a good deal of information about the code (see for example [1, §16.1] for some things you can do with the weight enumerator). But is has been calculated for only a few families of codes (e.g. Hamming codes [34], second order Reed Muller codes [41]).

OPEN PROBLEM 1. Find the weight enumerators of all Reed Muller codes (cf. [15]).

We mention in passing a related problem. The distribution of coset leaders by weight is also important for finding the error probability of a code, and for other reasons. But almost nothing is known about calculating it ([4],[12],[42]).

OPEN PROBLEM 2. Find the weight distribution of the coset leaders of the first order Reed Muller codes.

A code is *maximal self-orthogonal* if it is self-orthogonal and is not contained in any larger self-orthogonal code. For binary codes, a maximal self-orthogonal code has dimension $k = \frac{n-1}{2}$ if n is odd, or $k = \frac{n}{2}$ (and is self-dual) if n is even. This paper is concerned with weight enumerators of maximal self-orthogonal codes. First we give some examples.

EXAMPLES. These are binary codes

[n,k,d]	Code	Weight enumerator
1. [2,1,2]	{00,11}	$\phi_2 = x^2+y^2$.
2. [3,1,3]	{000,111}	$\phi_3 = x^3+y^3$.
3. [3,2,2]	{000,011,101,110}	x^3+3xy^2.
4. [7,3,4]	Even weight Hamming code	$\phi_7 = x^7+7x^3y^4$.
5. [7,4,3]	Hamming code	$x^7+7x^4y^3+7x^3y^4+y^7$.
6. [8,4,4]	Extended Hamming code	$\phi_8 = x^8+14x^4y^4+y^8$.
7. [17,8,4]	$I_{17}^{(3)}$ of [37]	$\phi_{17} = x^{17}+17x^{13}y^4+187x^9y^8+51x^5y^{12}$.
8. [23,11,8]	Even weight Golay code	$\phi_{23} = x^{23}+506x^{15}y^8+1288x^{11}y^{12}+$ $+253x^7y^{16}$.
9. [24,12,8]	Extended Golay code	$\phi_{24} = x^{24}+759(x^{16}y^8+x^8y^{16})+$ $+2576x^{12}y^{12}+y^{24}$.

(The subscript of a polynomial almost always gives its degree.)

All except examples 3 and 5 are maximal self-orthogonal. Examples 1, 6, 9 are self-dual.

Observe that examples 4, 6, 7, 8, 9 have the property that every codeword has weight divisible by 4 (because only powers of y^4 appear in the weight enumerators). Codes with this property are important because they have connections with block designs [1], sphere packings [17], lattices and finite groups [6-8], and projective planes [26] (see also [22]).

For non-binary codes it is often useful to have more detailed information than is given by the Hamming weight enumerator. Let C be a code over F_q, where the elements of F_q are labeled $\omega_0=0,\omega_1,\ldots,\omega_{q-1}$ in some fixed order. Then the *composition* of a vector $v \in F_q^n$ is defined to be $s(v) = (s_0(v),s_1(v),\ldots,s_{q-1}(v))$, where $s_i(v)$ denotes the number of coordinates of v that are equal to ω_i. Clearly $\sum_i s_i(v) = n$. Let $A(s)$ be the number of codewords $v \in C$ such that $s(v) = s$. Then the complete weight enumerator of C is the polynomial

$$W_C(z_0,\ldots,z_{q-1}) = \sum_s A(s) z_0^{s_0} \ldots z_{q-1}^{s_{q-1}}.$$

This is a homogeneous polynomial of degree n in the q indeterminates $z_0,\ldots,z_{q-1}$.

The next two examples are of self-dual codes over $F_3 = \{0,1,2\}$, having allweights divisible by 3. The exponents of x,y,z give the number of 0's,

1's, 2's respectively.

[n,k,d]	Code	Complete (V) and Hamming (W) weight enumerators
10. [4,2,3]	{0000,1110,0121, 2220,0212,1201, 1022,2011,2102}	$V = \Psi_4 = x^4+x(y+z)^3$, $W = \psi_4 = x^4+8xy^3$.
11. [12,6,6]	Extended Golay code over GF(3), containing the vector 11...1	$V = \Psi_{12} = x^{12}+y^{12}+z^{12}+22(x^6y^6+x^6z^6+y^6z^6)+$ $+220(x^6y^3z^3+x^3y^6z^3+x^3y^3z^6)$, $W = \psi_{12} = x^{12}+264x^6y^6+440x^3y^9+24y^{12}$.

2. MACWILLIAMS' THEOREM

This theorem, due to Mrs. F.J. MACWILLIAMS [20,21], is one of the most remarkable results in coding theory. It says that the weight enumerator of the dual code $C^\perp$ is completely determined by the weight enumerator of C.

We shall prove the binary case. The proof depends on the following lemma, which can be considered as a version of the Poisson summation formula [9,p.220]. Here $F = F_2$.

LEMMA 2.1. (cf. [19]). *Let* $f: F^n \to A$ *be any mapping from* F^n *into a vector space* A *over the complex numbers. Define the Fourier transform* $\hat{f}: F^n \to A$ *by*

$$\hat{f}(u) = \sum_{v \in F^n} f(v)(-1)^{u \cdot v}.$$

Then for any linear code $C \subset F^n$ *we have*

$$\sum_{v \in C^\perp} f(v) = \frac{1}{|C|} \sum_{u \in C} \hat{f}(u).$$

PROOF.

$$\sum_{u \in C} \hat{f}(u) = \sum_{u \in C} \sum_{v \in F^n} f(v)(-1)^{u \cdot v} = \sum_{v \in F^n} f(v) \sum_{u \in C} (-1)^{u \cdot v}.$$

If $v \in C^\perp$, the inner sum is equal to $|C|$. But if $v \notin C^\perp$, $u \cdot v = 0$ and 1 equally often and the inner sum is zero. □

THEOREM 2.1.(MACWILLIAMS' theorem, binary case). *Let C be an* [n,k] *binary code and* $C^{\perp}$ *its dual code. Then the weight enumerator of* $C^{\perp}$ *is given by*

$$W_{C^{\perp}}(x,y) = \frac{1}{2^k} W_C(x+y,x-y).$$

PROOF. In the lemma, let A be the set of polynomials in x,y with complex coefficients, and $f(v) = x^{n-wt(v)}y^{wt(v)}$. Then

$$\hat{f}(u) = \sum_{v\in F^n} x^{n-wt(v)}y^{wt(v)}(-1)^{u\cdot v} =$$

$$= \sum_{v_1=0}^{1}\sum_{v_2=0}^{1}\cdots\sum_{v_n=0}^{1}\prod_{i=1}^{n} x^{1-v_i}y^{v_i}(-1)^{u_iv_i} =$$

$$= \prod_{i=1}^{n}\sum_{v=0}^{1} x^{1-v}y^{v}(-1)^{u_iv}.$$

If $u_i = 0$ the inner sum is x+y; if $u_i = 1$ the inner sum is x-y. Therefore

$$\hat{f}(u) = (x+y)^{n-wt(u)}(x-y)^{wt(u)}. \quad \square$$

Examples of MACWILLIAMS' theorem

(i) $C = \underline{0}$, $W_C = x^n$; $C^{\perp} = F^n$, $W_{C^{\perp}} = (x+y)^n$.

(ii) $C = \{\underline{0},\underline{1}\}$, $W_C = x^n+y^n$; $C^{\perp}$ = {even weight vectors of length n}, $W_{C^{\perp}} = \frac{1}{2}\{(x+y)^n+(x-y)^n\}$.

(iii) Example 1 of section 1: $C = \{00,11\} = C^{\perp}$, $W_C = x^2+y^2 = \phi_2$.

(iv) Verify the theorem for the pairs of examples 2 and 3, and 4 and 5.

We state without proof two more general versions.

THEOREM 2.2. (MACWILLIAMS' theorem for Hamming weight enumerators). *Let C be an* [n,k] *code over* F_q. *Then*

$$W_{C^{\perp}}(x,y) = \frac{1}{q^k} W_C(x+(q-1)y,x-y).$$

For the next theorem we need a little more notation. Let $q = p^a$ where p is prime. Let f(x) be a primitive irreducible polynomial of degree a over F_p and let α be a root of f(x). Any element λ of F_q can be written uniquely as

$$\lambda = \lambda_0 + \lambda_1\alpha + \ldots + \lambda_{a-1}\alpha^{a-1}, \qquad 0 \le \lambda_i < p.$$

Let $\xi = e^{2\pi i/p}$, a complex p-th root of unity. Then the mapping $\chi: \lambda \to \xi^{\lambda_0}$ is a *character* of F_q, i.e., a homomorphism from the additive group of F_q to the multiplicative group of the complex numbers. E.g., if $q = p = 2$, χ maps $x \in F_2$ onto $(-1)^x$.

THEOREM 2.3. (MACWILLIAMS' theorem for complete weight enumerators.) *Let C be an* [n,k] *code over* F_q. *Then the complete weight enumerator of* $C^\perp$ *is given by*

$$V_{C^\perp}(z_0,\ldots,z_{q-1}) = \frac{1}{q^k} V_C\left(\sum_{j=0}^{q-1} \chi(\omega_0\omega_j)z_j, \ldots, \sum_{j=0}^{q-1} \chi(\omega_{q-1}\omega_q)z_j\right).$$

EXAMPLE. $q = 3$, $\xi = \omega = e^{2\pi i/3}$.

$$V_{C^\perp}(x,y,z) = \frac{1}{3^k} V_C(x+y+z, x+\omega y+\omega^2 z, x+\omega^2 y+\omega z).$$

Verify that theorems 2.2 and 2.3 hold for the code of example 10. For the proofs of theorems 2.2 and 2.3 see [20],[21],[19], and for other generalizations see [23],[24].

3. GLEASON'S THEOREM

If the code is self-dual, $C = C^\perp$, then the MACWILLIAMS' theorems 2.1, 2.2, 2.3 give identities which the weight enumerators must satisfy. For example theorem 2.1 states that the weight enumerator of a binary self-dual code must satisfy

$$W(x,y) = \frac{1}{2^{n/2}} W(x+y, x-y),$$

or since W(x,y) is homogeneous of degree n,

$$W(x,y) = W\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right). \tag{3.1}$$

From this GLEASON [11] was able to prove the following remarkable theorems.

THEOREM 3.1. (GLEASON's theorem I). *Let C be a binary self-dual code. Then the weight enumerator* $W_C(x,y)$ *of C is a polynomial in the weight enumerators*

$$\phi_2 = x^2 + y^2 \quad \textit{and} \quad \phi_8 = x^8 + 14x^4y^4 + y^8$$

of section 1. Equivalently, W *is a polynomial in* ϕ_2 *and*

$$\theta_8 = x^2y^2(x^2-y^2)^2 = \tfrac{1}{4}(\phi_2^4-\phi_8).$$

THEOREM 3.2. (GLEASON's theorem II). *Let C be a binary self-dual code in which every codeword has weight divisible by 4. Then the weight enumerator* $W_C(x,y)$ *of C is a polynomial in the weight enumerators* ϕ_8 *and*

$$\phi_{24} = x^{24} + 759x^{16}y^8 + 2576x^{12}y^{12} + 759x^8y^{16} + y^{24}$$

of section 1; or equivalently in ϕ_8 *and*

$$\theta_{24} = x^4y^4(x^4-y^4)^4 = \frac{1}{42}(\phi_8^3-\phi_{24}).$$

THEOREM 3.3. (GLEASON's theorem III). *Let C be a self-dual code over* F_3. *Then the Hamming weight enumerator* $W_C(x,y)$ *of C is a polynomial in the weight enumerators*

$$\psi_4 = x^4 + 8xy^3 \quad \textit{and} \quad \psi_{12} = x^{12} + 264x^6y^6 + 440x^3y^9 + 24y^{12}$$

of section 1; or equivalently in ψ_4 *and*

$$\theta_{12} = y^3(x^3-y^3)^3 = \frac{1}{24}(\psi_4^3-\psi_{12}).$$

Applications of theorem 3.1.

(i) A self-dual code of length 12 contains no codewords of weight 2. What is its weight enumerator W? Answer: By theorem 3.1 W has the form

$$W = a_1\phi_2^6 + a_2\phi_2^2\theta_8 =$$
$$= a_1(x^{12}+6x^{10}y^2+\ldots) + a_2(x^4+2x^2y^2+y^4)x^2y^2(x^2-y^2)^2.$$

But since there are no words of weight 2, this is also

$$= x^{12} + 0x^{10}y^2 + \ldots .$$

Therefore $a_1 = 1$, $a_2 = -6$, and

$$W = x^{12} + 15x^8y^4 + 32x^6y^6 + 15x^4y^8 + y^{12}.$$

(ii) Is there a self-dual code of length 32 with minimum distance 10?

Answer: By theorem 3.1 its weight enumerator W has the form

$$W = a_1\phi_2^{16} + a_2\phi_2^{12}\theta_8 + a_3\phi_2^8\theta_8^2 + a_4\phi_2^4\theta_8^3 + a_5\theta_8^4 =$$

$$= x^{32} + 0x^{30}y^2 + 0x^{28}y^4 + 0x^{26}y^6 + 0x^{24}y^8 + A_{10}x^{22}y^{10} + \ldots .$$

Equating coefficients we find that $a_1,\ldots,a_5$ are uniquely determined and that

$$W = x^{32} + 4960x^{22}y^{10} - 3472x^{20}y^{12} + \ldots .$$

Since a weight enumerator cannot have a negative coefficient, no such code exists.

(iii) The extended Golay code (example 9) has

$$W = \phi_{24} = \phi_2^{12} - 12\phi_2^8\theta_8 + 6\phi_2^4\theta_8^2 - 64\theta_8^3.$$

(iv) Exercise: Take all the codewords in the extended Golay code which begin either with 00... or 11..., and delete the first two coordinates. Use theorem 3.1 to obtain the weight distribution of this code. (Answer: $x^{22} + y^{22} + 77(x^{16}y^6+x^6y^{16}) + 330(x^{14}y^8+x^8y^{14}) + 616(x^{12}y^{10}+x^{10}y^{12})$).

Application of theorem 3.2.

The extended quadratic residue [48,24,12] code ([1,p.433]) has weights divisible by 4, so its weight enumerator has the form (from theorem 3.2)

(3.2) $$W = a_0\phi_8^6 + a_1\phi_8^3\theta_{24} + a_2\theta_{24}^2 =$$

$$= a_0(x^8+14x^4y^4+y^8)^6 + a_1x^4y^4(x^4-y^4)^4(x^8+14x^4y^4+y^8)^3 +$$

$$+ a_2x^8y^8(x^4-y^4)^8.$$

But since the minimum weight of this code is 12, W is also equal to

(3.3) $$x^{48} + 0x^{44}y^4 + 0x^{40}y^8 + \ldots .$$

Equating coefficients in (3.2), (3.3) we find a_0, a_1, a_2 are uniquely determined: $a_0 = 1$, $a_1 = -84$, $a_2 = 246$, and

(3.4) $$W = x^{48} + 17296x^{36}y^{12} + 535095x^{32}y^{16} + 3995376x^{28}y^{20} +$$

$$+ 7681680x^{24}y^{24} + 3995376x^{20}y^{28} + \ldots .$$

This example shows how powerful theorem 3.2 can be in obtaining weight enumerators: the fact that the minimum weight was 12 was enough to determine the full weight distribution!

We return to this example, and give further consequences of theorem 3.2, in section 6.

4. INVARIANT THEORY, AND PROOF OF GLEASON'S THEOREMS

Introduction

Other methods of proof are possible (see [3]) but the following proof from invariant theory is the simplest, once the necessary machinery has been developed, and is the easiest to generalize.

Suppose C is a binary self-dual code with weight enumerator $W(x,y)$. We have already seen in (3.1) that $W(x,y)$ must satisfy

(4.1) $$W(x,y) = W\left(\frac{x+y}{\sqrt{2}}, \frac{x-y}{\sqrt{2}}\right) .$$

Since C is self-dual, for any $x \in C$, $x \cdot x = 0$, so x has even weight, and only even powers of y appear in $W(x,y)$. Therefore

(4.2) $\qquad W(x,y) = W(x,-y).$

For an $n\times n$ matrix $A = (a_{ij})$ and a polynomial $f(\underline{x}) = f(x_1,\ldots,x_n)$, the result of transforming the variables of f by A is denoted $A\circ f(x) = f(\sum a_{1j}x_j,\ldots,\sum a_{nj}x_j)$. Note that $B\circ(A\circ f(\underline{x})) = (AB)\circ f(\underline{x})$.

So (4.1), (4.2) state that $T_1\circ W = W$, $T_2\circ W = W$, where

$$T_1 = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \qquad T_2 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}.$$

This implies $T\circ W = W$, where T is any matrix in the group generated by T_1 and T_2. We denote this group by $G_1 = \langle T_1, T_2\rangle$. It is not difficult to check that G_1 is isomorphic to the dihedral group of order 16.

Invariants

Let G be any finite group of $n\times n$ complex matrices, and let χ be a homomorphism from G into the multiplicative group of the complex numbers. (I.e., χ is a character of G.) Then $f(\underline{x})$ is called a *relative invariant of* G *with respect to* χ if

$$A\circ f(\underline{x}) = \chi(A)f(\underline{x}) \qquad \text{for all } A \in G.$$

In particular, if χ is identically 1, and

$$A\circ f(\underline{x}) = f(\underline{x}) \qquad \text{for all } A \in G,$$

then $f(\underline{x})$ is called an (*absolute*) *invariant* of G.

Clearly if f, g are absolute invariants of G so are $f+g$ and fg, so the absolute invariants form a ring. If f is an absolute invariant and g,h are relative invariants with respect to χ, then $g+h$ and fg are also relative invariants with respect to χ.

(4.1), (4.2) state that $W(x,y)$ is invariant under G_1. To prove theorem 3.1, it will be sufficient to specify the ring of invariants of G_1.

For any finite group G let $J(G)$ denote the ring of absolute invariants. The problem of characterizing $J(G)$ is very old, and there are many classical results (see [5,Chapter 17],[32,Part II], [44]). It is enough to characterize the invariants which are homogeneous polynomials, since any invariant is a sum of homogeneous invariants.

Existence of a basic set of invariants for finite groups

DEFINITION 4.1. Polynomials $f_1(\underline{x}),\ldots,f_m(\underline{x})$ are *algebraically dependent* if there is a polynomial p with complex coefficients, not all zero, such that $p(f_1(\underline{x}),\ldots,f_m(\underline{x}))$ is identically zero. Otherwise $f_1(\underline{x}),\ldots,f_m(\underline{x})$ are *algebraically independent.*

THEOREM 4.1. ([13,p.154]). *Any* n+1 *polynomials in* n *variables are algebraically dependent.*

By far the most convenient description of $J(G)$ is a set of invariants $f_1,\ldots,f_m$ such that any invariant is a *polynomial* in $f_1,\ldots,f_m$. Then $f_1,\ldots,f_m$ is called a *polynomial basis* for $J(G)$. By theorem 4.1 if $m > n$ there will be polynomial equations, which are called *syzygies*, relating $f_1,\ldots,f_m$.

E. NOETHER'S THEOREM 4.2. (cf. [44,pp.275-276]). $J(G)$ *has a polynomial basis consisting of not more than* $\binom{g+n}{n}$ *invariants, of degree not exceeding* g, *where* g *is the order of* G.

Theorem 4.2 says that a polynomial basis for $J(G)$ can always be found. Finding invariants is fairly easy using:

THEOREM 4.3. *If* $f(x)$ *is any polynomial then*

$$h(\underline{x}) = \sum_{A\in G} A\circ f(x)$$

is an invariant of G.

PROOF. For any $A' \in G$,

$$A'\circ h(\underline{x}) = \sum_{A\in G} A'\circ(A\circ f(\underline{x})) = \sum_{A\in G} (AA')\circ f(\underline{x}) =$$

$$= \sum_{B\in G} B\circ f(\underline{x}) = h(\underline{x}). \quad \square$$

Furthermore, it is clear that all invariants of G can be obtained in this way. In fact the proof of theorem 4.2 shows that a polynomial basis for the invariants of G can be obtained by averaging over G all monomials

$$x_1^{b_1} x_2^{b_2} \ldots x_n^{b_n}$$

of total degree $\sum b_i \leq g$.

More generally, any symmetric function of the g polynomials $\{A \circ f(\underline{x}) \mid A \in G\}$ is an invariant of G.

MOLIEN's theorem

The next three theorems enable one to determine when enough invariants have been found to make a basis:

THEOREM 4.4. (cf. [32,p.258]). *The number of linearly independent invariants of* G *of the first degree is*

$$\frac{1}{|G|} \sum_{A \in G} \text{trace}(A).$$

THEOREM 4.5. (MOLIEN [33],[32,p.259]). *The number of linearly independent invariants of* G *of degree* ν *is the coefficient of* λ^ν *in the expansion of*

$$\Phi(\lambda) = \frac{1}{|G|} \sum_{A \in G} \frac{1}{\det|I-\lambda A|} . \tag{4.3}$$

$\Phi(\lambda)$ is called the *Molien series* of G.

A similar result holds for relative invariants:

THEOREM 4.6. (MOLIEN [33],[32,p.259]). *The number of linearly independent relative invariants with respect to* χ *of degree* ν *is the coefficient of* λ^ν *in the expansion of*

$$\frac{1}{|G|} \sum_{A \in G} \frac{\bar{\chi}(A)}{\det|I-\lambda A|} , \tag{4.4}$$

where the bar denotes the complex conjugate.

A simple example

Let C be a self-dual code over GF(q) with Hamming weight enumerator W(x,y). By theorem 2.2, W(x,y) is invariant under the transformation

$$T_3 = \frac{1}{\sqrt{q}} \begin{pmatrix} 1 & q-1 \\ 1 & -1 \end{pmatrix} .$$

Now $T_3^2 = I$, so T_3 generates the group $G_2 = \{I, T_3\}$ of order 2. We shall find the invariants of G_2.

By averaging x over the group, using theorem 4.3, we obtain the invariant $x + \frac{1}{\sqrt{q}}(x+(q-1)y)$, or equivalently $\sigma_1 = x + (\sqrt{q}-1)y$. By averaging x^2 we obtain the invariant $x^2 + \frac{1}{q}(x+(q-1)y)^2$, or equivalently, subtracting $(1+1/q)\sigma_1^2$, $\sigma_2 = y(x-y)$.

Any polynomial in σ_1,σ_2 is of course an invariant of G_2, and the number of products $\sigma_1^i\sigma_2^j$ of degree ν is equal to the number of solutions of $i+2j = \nu$, which is the coefficient of λ^ν in

$$(4.5) \qquad (1+\lambda+\lambda^2+\ldots)(1+\lambda^2+\lambda^4+\ldots) = \frac{1}{(1-\lambda)(1-\lambda^2)} .$$

To see if this includes all the invariants of G_2 we compute the Molien series (4.3). This is

$$\Phi(\lambda) = \tfrac{1}{2}\left(\frac{1}{(1-\lambda)^2} + \frac{1}{1-\lambda^2}\right) = \frac{1}{(1-\lambda)(1-\lambda^2)}$$

which agrees with (4.5)! We conclude that we have found all the invariants, i.e., that σ_1,σ_2 are a polynomial basis for the invariants of G_2.

For coding applications we are interested in invariants of even degree. This corresponds to extending G_2 by adding the matrix -I, and the Molien series becomes

$$\Phi_e(\lambda) = \tfrac{1}{2}\ (\Phi(\lambda)+\Phi(-\lambda)) = \frac{1}{(1-\lambda^2)^2} ,$$

and as a basis for the new invariants we may take σ_1^2,σ_2^2 or equivalently $\sigma_3 = x^2+(q-1)xy$, $\sigma_4 = xy-y^2$. Thus we have shown that the Hamming weight enumerator of any self-dual code over GF(q) is a polynomial in σ_3 and σ_4.

For example, the code generated by {11} (which is self-dual if q is even) has weight enumerator $\sigma_3-(q-1)\sigma_4$.

The preceding argument enables us to give a short proof of a recent result of LEONT'EV.

THEOREM 4.7. (LEONT'EV [18]). *For a linear code C over* GF(q)

$$W_C(x,y)W_C\left(\frac{x+(q-1)y}{\sqrt{q}},\frac{x-y}{\sqrt{q}}\right)$$

is a polynomial in σ_3 *and* σ_4.

PROOF. This product is clearly invariant under T_3 and $-I$, and so the result follows from what we have just proved. □

Notation. The following notation is convenient for describing invariants. $\mathbb{C}$ denotes the complex numbers. If f,g,h,... are polynomials $\mathbb{C}[f,g,h,\ldots]$ denotes the ring of polynomials in f,g,h,... with complex coefficients. If R and S are rings, $R \oplus S$ denotes their *direct sum*.

Using this notation, we see that the following result implies theorem 3.1.

THEOREM 3.1*. *The ring of invariants of* $G_1 = \langle T_1, T_2 \rangle$ *is* $\mathbb{C}[\phi_2, \phi_8]$.

PROOF. Let $J(G_1)$ denote the ring of invariants of G_1. We know from coding theory that ϕ_2 and ϕ_8 are in $J(G_1)$, and so $J(G_1) \supseteq M = \mathbb{C}[\phi_2, \phi_8]$.

To show $J(G_1) = M$, let a_d (or b_d) be the number of linearly independent polynomials of degree d in $J(G_1)$ (or M). Clearly

$$\sum_{d=0}^{\infty} b_d \lambda^d = \frac{1}{(1-\lambda^2)(1-\lambda^8)} .$$

But from MOLIEN's theorem 4.5,

$$\Phi(\lambda) = \sum_{d=0}^{\infty} a_d \lambda^d = \frac{1}{16} \sum_{A \in G_1} \frac{1}{|I-\lambda A|} = \frac{1}{(1-\lambda^2)(1-\lambda^8)} \tag{4.6}$$

after a straightforward calculation. Therefore $a_d = b_d$ for all d, and so $J(G_1) = M$. □

In a similar manner we deduce theorems 3.2 and 3.3 from:

THEOREM 3.2*. *The ring of invariants of the group* $G_3 = \langle T_1, T_4 \rangle$, *where*

$$T_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T_4 = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix},$$

is $\mathbb{C}[\phi_8, \phi_{24}]$.

G_3 has order 192, and the Molien series (4.3) is

$$\Phi(\lambda) = \frac{1}{(1-\lambda^8)(1-\lambda^{24})} . \tag{4.7}$$

THEOREM 3.3*. *The ring of invariants of* $G_4 = \langle T_5, T_6 \rangle$, *where*

$$T_5 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 2 \\ 1 & -1 \end{pmatrix}, \qquad T_6 = \begin{pmatrix} 1 & 0 \\ 0 & e^{2\pi i/3} \end{pmatrix}.$$

is $C[\psi_4, \psi_{12}]$.

G_4 has order 48, and the Molien series (4.3) is

$$\Phi(\lambda) = \frac{1}{(1-\lambda^4)(1-\lambda^{12})}. \tag{4.8}$$

Theorems 3.1*, 3.2*, 3.3* must have been known to KLEIN and BURNSIDE (see [16],[5,p.362],[39]).

Note that in all these examples the degrees of the basic invariants can be read off the Molien series (4.5)-(4.8) (cf. [28]).

5. GENERALIZATIONS OF GLEASON'S THEOREMS

In this section we give three generalizations of GLEASON's theorems. Other generalizations will be found in [23],[29]. The proofs always follow the same procedure:

Step 1. Translate assumptions about the code into algebraic constraints on the weight enumerator.

Step 2. Use invariant theory to find all possible polynomials satisfying these constraints.

But each of the three examples given has special features of its own. Theorem 5.1 is unusual in that it is rather difficult to find the basic invariants. (Usually one quickly finds what one thinks is a basis for the invariants and the difficulty lies in proving that it *is* a basis.) Theorems 5.2 and 5.3 use a group whose order becomes arbitrarily large, and theorem 5.4 also requires the introduction of new indeterminates and the use of relative rather than absolute invariants.

(I) COMPLETE WEIGHT ENUMERATOR OF A TERNARY SELF-DUAL CODE

Let C be an $[n, \frac{1}{2}n]$ self-dual code over GF(3) which contains the codeword $\underline{1} = 11\ldots1$. Let the complete weight enumerator of C be

$$V(x,y,z) = \sum_{u \in C} x^{s_0(u)} y^{s_1(u)} z^{s_2(u)},$$

where $s_i(u)$ is the number of components of u which are equal to i (as in section 1).

THEOREM 5.1. (cf. [30]).

$$V(x,y,z) \in \mathbb{C}[\alpha_{12},\beta_{12},\delta_{36}] \oplus \gamma_{24}\mathbb{C}[\alpha_{12},\beta_{12},\delta_{36}]$$

(i.e., V(x,y,z) can be written uniquely as a polynomial in α_{12}, β_{12}, δ_{36} *plus* γ_{24} *times another such polynomial), where*

$$a = x^3 + y^3 + z^3,$$

$$p = 3xyz,$$

$$b = x^3y^3 + x^3z^3 + y^3z^3,$$

$$\alpha_{12} = a(a^3+8p^3),$$

$$\beta_{12} = (a^2-12b)^2,$$

$$\gamma_{24} = b[(9b-a^2)^3 - 3ap^3(9b-a^2) - a^3p^3 - p^6],$$

$$\delta_{36} = p^3(a^3-p^3)^3.$$

(As usual, the subscript of a polynomial gives its degree.)

PROOF. We carry out the two steps just mentioned.

Step 1. Let a typical codeword $u \in C$ contain a 0's, b 1's, and c 2's. Then since C is self-dual and contains $\underline{1}$

$$u\cdot u = 0 \pmod 3 \Rightarrow 3 \mid b + c,$$

$$u\cdot \underline{1} = 0 \pmod 3 \Rightarrow 3 \mid b - c \Rightarrow 3 \mid b \text{ and } 3 \mid c,$$

$$\underline{1}\cdot\underline{1} = 0 \pmod 3 \Rightarrow 3 \mid a + b + c \Rightarrow 3 \mid a.$$

Therefore V(x,y,z) is invariant under the transformations

$$\begin{pmatrix} \omega & & \\ & 1 & \\ & & 1 \end{pmatrix}, \quad J = \begin{pmatrix} 1 & & \\ & \omega & \\ & & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & & \\ & 1 & \\ & & \omega \end{pmatrix}, \quad \omega = e^{2\pi i/3}.$$

Also -u contains a 0's, c 1's, b 2's, and $\underline{1}$ + u contains c 0's, a 1's, b 2's. Therefore V(x,y,z) is invariant under

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

i.e., under any permutation of its arguments.

Finally from the MACWILLIAMS' theorem (the example after theorem 2.3), V(x,y,z) is invariant under

$$M = \frac{1}{\sqrt{3}}\begin{pmatrix}1&1&1\\1&\omega&\omega^2\\1&\omega^2&\omega\end{pmatrix}.$$

These 6 matrices generate a group G_5, of order 2592, consisting of 1944 matrices of the type

$$s^\nu\begin{pmatrix}1&&\\&\omega^a&\\&&\omega^b\end{pmatrix} M^e\begin{pmatrix}1&&\\&\omega^c&\\&&\omega^d\end{pmatrix}, \quad s = e^{2\pi i/12},$$

and 648 matrices of the type

$$s^\nu\begin{pmatrix}1&&\\&\omega^a&\\&&\omega^b\end{pmatrix} P,$$

where $0 \le \nu \le 11$, $0 \le a,b,c,d \le 2$, $e = 1$ or 3, and P is any 3×3 permutation matrix.

Step 2. This step consists of showing that the ring of invariants of G_5 is equal to $\mathbb{C}[\alpha_{12},\beta_{12},\delta_{36}] \oplus \gamma_{24}\mathbb{C}[\alpha_{12},\beta_{12},\delta_{36}]$. First, it is straightforward to show that the Molien series (4.3) of G_5 is

$$\Phi(\lambda) = \frac{1+\lambda^{24}}{(1-\lambda^{12})^2(1-\lambda^{36})},$$

which suggests the degrees of the basic invariants that we should look for.

Next, G_5 is generated by J, M, and all permutation matrices P. Obviously the invariants must be symmetric functions of x,y,z. So we take the symmetric functions a,p,b, and find functions of them which are invariant under J and M. For example,

$$a \overset{M}{\longleftrightarrow} \frac{1}{\sqrt{3}}(a+2p) \overset{J}{\longrightarrow} \frac{1}{\sqrt{3}}(a+2\omega p) \overset{M}{\longleftrightarrow} \frac{i}{\sqrt{3}}(a+2\omega^2 p),$$

so one invariant is

$$\alpha_{12} = a(a+2p)(a+2\omega p)(a+2\omega^2 p) = a(a^3+8p^3).$$

Again

$$a^2-12b \overset{M}{\longleftrightarrow} -(a^2-12b),$$

so another invariant is $\beta_{12} = (a^2-12b)^2$. Again

$$b \overset{M}{\longleftrightarrow} \frac{1}{9}(a^2+ap+p^2) - b \overset{J}{\longrightarrow} \frac{1}{9}(a^2+\omega ap+\omega^2 p^2) - b \overset{M}{\longleftrightarrow}$$

$$- \frac{1}{9}[(a^2+\omega^2 ap+\omega p^2)-b]$$

gives the invariant

$$\gamma_{24} = b(9b-a^2-ap-p^2)(9b-a^2-\omega ap-\omega^2 p^2)(9b-a^2-\omega^2 ap-\omega p^2) =$$

$$= b[(9b-a^2)^3-a^3p^3-p^6-3ap^3(9b-a^2)].$$

Finally

$$p \overset{M}{\longleftrightarrow} \frac{1}{\sqrt{3}}(a-p) \overset{J}{\longrightarrow} \frac{1}{\sqrt{3}}(a-\omega p) \overset{M}{\longleftrightarrow} \frac{s}{\sqrt{3}}(a-\omega^2 p)$$

gives the invariant

$$\delta_{36} = p^3(a-p)^3(a-\omega p)^3(a-\omega^2 p)^3 = p^3(a^3-p^3)^3.$$

One can show that $\alpha_{12}, \beta_{12}, \delta_{36}$ are algebraically independent and that there is a syzygy of degree 48:

$$(768\gamma_{24}+\alpha_{12}^2+18\alpha_{12}\beta_{12}-27\beta_{12}^2)^2 = 64\beta_{12}(\alpha_{12}^3-64\delta_{36}). \quad \square$$

Remark. Without the assumption that the code contains the all-ones vector the theorem (due to R.J. McEliece [23,§4.7]) becomes much more complicated.

Applications of theorem 5.1.

For the ternary Golay code (example 11 of section 1), $V = \alpha_{12} + \frac{4}{3}\beta_{12}$.
For PLESS's [24,12,9] symmetry code ([35],[36]),

$$V = \frac{179}{432}\alpha_{12}^2 - \frac{19}{24}\alpha_{12}\beta_{12} + \frac{595}{432}\beta_{12}^2 - \frac{352}{9}\gamma_{24}.$$

We have also found the complete weight enumerator of the symmetry codes of lengths 36 and 48 ([30]).

(II) SPLIT WEIGHT ENUMERATORS

We define the *left* and *right weight* of a vector $v = (v_1,\ldots,v_m,v_{m+1},\ldots,v_{2m})$ to be respectively $w_L = wt(v_1,\ldots,v_m)$ and $w_R = wt(v_{m+1},\ldots,v_{2m})$. The *split weight enumerator* of a [2m,k] binary code C is

$$W_C(x,y,X,Y) = \sum_{v\in C} x^{m-w_L(v)} y^{w_L(v)} X^{m-w_R(v)} Y^{w_R(v)}.$$

THEOREM 5.2. (cf. [29]). *Let C be a* [2m,m] *self-dual binary code satisfying:*

(B1) *C contains the vectors* $0^m1^m = 0\ldots01\ldots1$ *and* $\underline{1}$; *and*

(B2) *the number of codewords with* $(w_L,w_R) = (j,k)$ *is equal to the number with* $(w_L,w_R) = (k,j)$.

Then

(i) $W_C(x,y,X,Y)$ *is an element of* $\mathbb{C}[\rho_4,\eta_8,\theta_{16}]$, *where*

$$\rho_4 = (x^2+y^2)(X^2+Y^2),$$
$$\eta_8 = x^4X^4 + x^4Y^4 + y^4X^4 + y^4Y^4 + 12x^2y^2X^2Y^2,$$
$$\theta_{16} = (x^2X^2-y^2Y^2)^2(x^2Y^2-y^2X^2)^2.$$

(ii) *Furthermore, if all weights in C are multiples of 4, then* $W_C(x,y,X,Y)$ *is an element of* $\mathbb{C}[\eta_8,\theta_{16},\gamma_{24}]$, where

$$\gamma_{24} = x^2y^2X^2Y^2(x^4-y^4)^2(X^4-Y^4)^2.$$

A code satisfying (B1), (B2) *is "balanced" about its midpoint, and the division into two halves is a natural one.*

Applications of theorem 5.2.

If $u = (u_1,\ldots,u_n)$ and $v = (v_1,\ldots,v_n)$ let $u|v = (u_1,\ldots,u_n,v_1,\ldots,v_n)$. For $j = 1,2$ let C_j be a code of length n with ordinary weight enumerator $W_j(x,y)$ and split weight enumerator $W_j(x,y,X,Y)$. The code $C_1|C_2 = \{u|v: u \in C_1,\ v \in C_2\}$ has ordinary and split weight enumerators $W_1(x,y)W_2(x,y)$ and $W_1(x,y)W_2(X,Y)$. The equivalent code $C_1\|C_2 = \{u'|v'|u''|v'':$ $: u=u'|u'' \in C_1;\ v=v'|v'' \in C_2\}$, where u and v are broken in half, has ordi-

nary and split weight enumerators $W_1(x,y)W_2(x,y)$ and $W_1(x,y,X,Y)W_2(x,y,X,Y)$.

There is a MACWILLIAMS theorem for split weight enumerators (easily obtained from lemma 2.1):

$$(5.1) \qquad W_{C^\perp}(x,y,X,Y) = \frac{1}{|C|} W_C(x+y,x-y,X+Y,X-Y).$$

We use a detached-coefficient notation for W, and instead of the terms

$$\alpha(x^a y^b X^c Y^d + x^a y^b X^d Y^c + x^b y^a X^c Y^d + x^b y^a X^d Y^c)$$

we write a row of a table:

c/0	x	y	X	Y	#
α	a	b	c	d	4

giving respectively the coefficient, the exponents, and the number of terms of this type. The sum of the products of the first and last columns is the total number of codewords.

A quadratic residue code of length 81 = q+1, where q is a prime, with

Table I. Split weight enumerators

Code	W	c/0	x	y	X	Y	#
H_8	η_8	1	4	0	4	0	4
		12	2	2	2	2	1
	θ_{16}	1	8	0	4	4	4
		-2	6	2	6	2	4
		4	4	4	4	4	1
	γ_{24}	1	10	2	10	2	4
		-2	10	2	6	6	4
		4	6	6	6	6	1
G_{24}		1	12	0	12	0	4
		132	10	2	6	6	4
		495	8	4	8	4	4
		1584	6	6	6	6	1
Z_{48}		1	24	0	24	0	4
		276	22	2	14	10	8
		3864	20	4	16	8	8
		13524	20	4	12	12	4
		9016	18	6	18	6	4
		125580	18	6	14	10	8
		256335	16	8	16	8	4
		950544	16	8	12	12	4
		1835400	14	10	14	10	4
		3480176	12	12	12	12	1

generator matrix, in the canonical form of figures 1-7 of [14], satisfies the hypotheses of theorem 5.2(ii). Table I shows 3 such codes, the [8,4,4] Hamming code H_8 (example 6 of section 1), the [24,12,8] extended Golay code G_{24} (example 9 of section 1) for which $W = \eta_8^3 - 3\eta_8\theta_{16} - 42\gamma_{24}$, and the [48,24,12] quadratic residue code Z_{48}. Also if $S_2 = \{00,11\}$, $S_2|S_2$ has $W = \rho_4$. $H_8|H_8$ has $W = \eta_8^2 + 12\theta_{16}$. Let $R(r,m)$ denote an r-th order Reed-Muller (RM) code of length 2^m. Then RM codes can be constructed recursively from $R(r+1,m) * R(r,m) = R(r+1,m+1)$, where $C_1 * C_2 = \{u|(u+v) : u \in C_1, v \in C\}$ [43]. The first order RM code of length n obtained in this way has

$$W = (x^{\frac{1}{2}n}+y^{\frac{1}{2}n})(X^{\frac{1}{2}n}+Y^{\frac{1}{2}n}) + (2n-4)(xyXY)^{\frac{1}{4}n}.$$

PROOF OF THEOREM 5.2(ii). (Part (i) is similar.) Let C satisfy the hypotheses of theorem 5.2(ii) and have split weight enumerator $W = W(x,y,X,Y)$. From the hypotheses, equation (5.1), and the fact that in each term $x^j y^k X^l Y^m$ of W, $j+k = l+m$, it follows that W is invariant under

$$\frac{1}{\sqrt{2}}\begin{pmatrix}1&1&0&0\\1&-1&0&0\\0&0&1&1\\0&0&1&-1\end{pmatrix}, \begin{pmatrix}1&&&\\&i&&\\&&1&\\&&&i\end{pmatrix}, \begin{pmatrix}1&&&\\&1&&\\&&1&\\&&&1\end{pmatrix}, \begin{pmatrix}&&1&\\&&&1\\1&&&\\&1&&\end{pmatrix} \text{ and } \begin{pmatrix}\alpha&&&\\&\alpha&&\\&&\alpha^{-1}&\\&&&\alpha^{-1}\end{pmatrix},$$

for any complex number α. Let us choose α to be a primitive complex p-th root of unity, where p is a prime greater than deg W = length of C = 2m.

Then one can show that these matrices generate a group G_6 of order 6144p, and that the Molien series for G_6 is

$$(5.2) \qquad \frac{1}{(1-\lambda^8)(1-\lambda^{16})(1-\lambda^{24})} + O(\lambda^p).$$

(See [29] for details.) On the other hand, we know from coding theory that $J(G_6)$ contains $M = \mathbb{C}[\eta_8,\theta_{16},\gamma_{24}]$. But the number of linearly independent polynomials of degree 2m in M is the coefficient of λ^{2m} in

$$(5.3) \qquad \frac{1}{(1-\lambda^8)(1-\lambda^{16})(1-\lambda^{24})}.$$

Because p > 2m, the coefficients of λ^{2m} in (5.2) and (5.3) agree, and so $J(G_6) = M$. □

(III) WEIGHT ENUMERATORS OF MAXIMAL BINARY SELF-ORTHOGONAL CODES OF ODD LENGTH

THEOREM 5.3.

(A) (cf. [29]). *For* n *odd, let* C *be an* $[n,\frac{1}{2}(n-1)]$ *binary self-orthogonal code. Thus* $C^{\perp} = C \cup (\underline{1}+C)$. *Then*

(i) $W_C(x,y)$ *is an element of the direct sum* $x\mathbb{C}[\phi_2,\phi_8] \oplus \phi_7\mathbb{C}[\phi_2,\phi_8]$, *where* $\phi_2 = x^2+y^2$, $\phi_7 = x^7+7x^3y^4$, $\phi_8 = x^8+14x^4y^4+y^8$. *In words:* $W_C(x,y)$ *can be written in a unique way as* x *times a polynomial in* ϕ_2 *and* ϕ_8, *plus* ϕ_7 *times another such polynomial.*

(B) *Suppose in addition that all weights in* C *are multiples of* 4. *Then*

(ii) n *must be of the form* $8m\pm1$.

(iii) *If* $n = 8m-1$, *then* $W_C(x,y)$ *is an element of* $\phi_7\mathbb{C}[\phi_8,\phi_{24}] \oplus \phi_{23}\mathbb{C}[\phi_8,\phi_{24}]$, *where* $\phi_{23} = x^{23} + 506x^{15}y^8 + 1288x^{11}y^{12} + 253x^7y^{16}$, $\phi_{24} = x^4y^4(x^4-y^4)^4$.

(iv) *If* $n = 8m+1$, *then* $W_C(x,y)$ *is an element of* $x\mathbb{C}[\phi_8,\phi_{24}] \oplus \phi_{17}\mathbb{C}[\phi_8,\phi_{24}]$, *where* $\phi_{17} = x^{17} + 17x^{13}y^4 + 187x^9y^8 + 51x^5y^{12}$.

See the examples in section 1. Some other examples: The [31,15,8] quadratic residue code: $W = -14\phi_7\phi_{24} + \phi_{23}\phi_8$. The [47,23,12] QR code: $W = \frac{1}{7}\{-253\phi_7\phi_8^2\phi_{24} + \phi_{23}(7\phi_8^3-41\phi_{24})\}$. See [37] for additional examples.

It is not presently known if a projective plane of order 10 exists. If it does exist, then from [26] the rows of its incidence matrix generate a [111,55,12] code with

$$W = \frac{1}{7}\{\phi_7(-253\phi_8^{10}\phi_{24} + 24123\phi_8^7\phi_{24}^2 - 430551\phi_8^4\phi_{24}^3 + c_1\phi_8\phi_{24}^4) + \\ + \phi_{23}(7\phi_8^{11} - 825\phi_8^8\phi_{24} + 22077\phi_8^5\phi_{24}^2 + c_2\phi_8^2\phi_{24}^3)\},$$

where c_1, c_2 are constants, at present unknown.

PROOF OF THEOREM 5.3. Let C be a code of length 4m-1 satisfying the hypotheses A,B of theorem 5.3, with weight enumerator $W(\underline{x}) = W(x,y)$. Let

$$M = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}, \quad J = \begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix}, \quad R = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} = MJ^2M.$$

By the MACWILLIAMS theorem 2.1, $M\circ W(\underline{x}) = 2^{-1/2}(W(\underline{x}) + R\circ W(\underline{x}))$. Also $J\circ W(\underline{x}) = W(\underline{x})$. Let $\mathcal{M}$ be the set of all polynomials satisfying these two equations.

From coding theory M contains $N = \phi_7\mathbb{C}[\phi_8,\phi_{24}] \oplus \phi_{23}\mathbb{C}[\phi_8,\phi_{24}]$. To show $M = N$, let a_d (or b_d) be the number of linearly independent polynomials of degree d in M (or N). Clearly $\sum_{d=0}^{\infty} b_d\lambda^d = (\lambda^7+\lambda^{23})/(1-\lambda^8)(1-\lambda^{24})$. We show $M = N$ by showing $a_d = b_d$ for all d.

The key device is to consider not W(x,y) but f(u,v,x,y) = uW(x,y) + + vW(y,x). Then f(u,v,x,y) is invariant under

$$M^+ = \begin{pmatrix} M & 0 \\ 0 & M \end{pmatrix} \quad \text{and} \quad J^+ = \begin{pmatrix} J & 0 \\ 0 & J \end{pmatrix} \quad \text{acting on} \begin{pmatrix} u \\ v \\ x \\ y \end{pmatrix}.$$

As in theorem 5.2, let ω be a primitive complex p-th root of unity, where p is a prime greater than deg W = length of C. Then f(u,v,x,y) is a relative invariant under P = diag($\omega,\omega,1,1$) with respect to $\chi(P) = \omega$.

Now M,J generate a group G_{192} of order 192, consisting of the matrices

$$(5.4) \qquad r^\nu \begin{pmatrix} 1 & 0 \\ 0 & \alpha \end{pmatrix},\ r^\nu \begin{pmatrix} 0 & 1 \\ \alpha & 0 \end{pmatrix},\ \frac{r^\nu}{\sqrt{2}} \begin{pmatrix} 1 & \beta \\ \alpha & -\alpha\beta \end{pmatrix},$$

where $r = (1+i)/\sqrt{2}$, $0 \le \nu \le 7$, $\alpha,\beta \in \{1,i,-1,-i\}$ (cf. [23]). So M^+, J^+, P generate a group G of order 192p consisting of the matrices $\begin{pmatrix} \omega^\nu A & \\ & A \end{pmatrix}$, $0 \le \nu \le p-1$, $A \in G_{192}$. Then the set M^+ of the relative invariants of G with respect to $\chi(M^+) = \chi(J^+) = 1$, $\chi(P) = \omega$ is in 1-1-correspondence with M up to degree p-1. Therefore from theorem 5.1, for all $p > d$, a_d is the coefficient of λ^{d+1} in

$$\frac{1}{192p}\sum_{B\in G}\frac{\bar{\chi}(B)}{|I-\lambda B|} = \frac{1}{192}\sum_{A\in G_{192}}\frac{1}{p}\sum_{\nu=0}^{p-1}\frac{\omega^{-\nu}}{|I-\lambda A||I-\lambda\omega^\nu A|} \to$$

$$(\text{as } p\to\infty, |\lambda|<1) \quad \to \frac{1}{192}\sum_{A\in G_{192}}\frac{1}{|I-\lambda A|}\frac{1}{2\pi}\int_0^{2\pi}\frac{e^{-i\theta}d\theta}{|I-\lambda e^{i\theta}A|} =$$

$$= \frac{\lambda}{192}\sum_{A\in G_{192}}\frac{\text{trace}(A)}{|I-\lambda A|} \overset{(5.4)}{=} \lambda\frac{\lambda^7+\lambda^{23}}{(1-\lambda^8)(1-\lambda^{24})}.$$

This proves (iii) and half of (ii). The case n = 4m+1 is treated similarly, taking $M^+ = \begin{pmatrix} \bar{M} & 0 \\ 0 & M \end{pmatrix}$, $J^+ = \begin{pmatrix} \bar{J} & 0 \\ 0 & J \end{pmatrix}$. For part (i) we take $J = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$, obtaining a group of order 16p. □

6. VERY GOOD SELF-DUAL CODES DO NOT EXIST

Let C be a binary self-dual code with all weights divisible by 4 and of length $n = 8j = 24\mu+8\nu$, $\nu = 0, 1$ or 2. From GLEASON's theorem 3.1, the weight enumerator of C has the form

$$W = \sum_{i=0}^{\mu} a_i f^{j-3i} g^i, \tag{6.1}$$

where $f = 1+14y+y^2$, $g = y(1-y)^4$. (We have replaced x by 1 and y^4 by y.)

Suppose the $\mu+1$ coefficients a_i are chosen so that

$$W = W^* = 1 + A_{4\mu+4}y^{\mu+1} + A_{4\mu+8}y^{\mu+2} + \ldots . \tag{6.2}$$

This determines the a_i and A_i uniquely. The resulting W^* is the weight enumerator of that self-dual code with the greatest minimum weight we could hope to attain, and is called an *extremal* weight enumerator.

If a code exists with weight enumerator W^*, it has minimum weight $d^* = 4\mu+4$, unless $A_{4\mu+4}$ is accidentally zero, in which case $d^* \geq 4\mu+8$.

But it can be shown (cf. [27]) that $A_{4\mu+4}$, the number of codewords of minimum non-zero weight enumerator, is equal to:

$$\binom{n}{5}\binom{5\mu-2}{\mu-1}\Big/\binom{4\mu+4}{5}, \qquad \text{if } n = 24\mu,$$

$$\frac{1}{4}n(n-1)(n-2)(n-4)\frac{(5\mu)!}{\mu!(4\mu+4)!}, \qquad \text{if } n = 24\mu+8,$$

$$\frac{3}{2}n(n-2)\frac{(5\mu+2)!}{\mu!(4\mu+4)!}, \qquad \text{if } n = 24\mu+16,$$

and is never zero. This proves

THEOREM 6.1. (cf. [27]). *The minimum weight of a binary self-dual code of length* n *with all weights divisible by* 4 *is* $\leq 4[\frac{n}{24}] + 4$.

However, the next coefficient, $A_{4\mu+8}$, turns out to be negative if n is large (above about 3712), and so a self-dual code with weight enumerator W^* does not exist for large n. In fact one can show that no self-dual code can even have minimum distance within a constant of $\frac{n}{6}$, if n is sufficiently large:

THEOREM 6.2. (cf. [31]). *Let* b *be any constant. Suppose the* a_i *in* (6.1) *are chosen so that*

$$W = 1 + A_{4d}y^d + A_{4d+4}y^{d+1} + \dots ,$$

where $d \geq \frac{n}{6} - b$. *Then one of the coefficients* A_i *is negative, for all sufficiently large* n. *So a binary self-dual code of length* n, *weights divisible by* 4, *and minimum weight* d *does not exist for all sufficiently large* n.

On the other hand it is known that self-dual codes exist which meet the Gilbert bound ([25],[38]).

Similar results hold for self-dual codes over GF(3) (see [31]).

OPEN PROBLEM 3. What is the greatest n for which equality holds in theorem 6.1? (cf. [40],[10]).

ACKNOWLEDGEMENTS

This paper is based on joint work of F.J. MACWILLIAMS, C.L. MALLOWS, A.M. ODLYZKO, and the author; see [23],[27],[29] and [31].

REFERENCES

[1] ASMUS Jr., A.F. & H.F. MATTSON Jr., *New 5-designs*, J. Combinatorial Theory, 6 (1969) 122-151.

[2] BERLEKAMP, E.R., *Algebraic coding theory*, McGraw-Hill, New York, 1968.

[3] BERLEKAMP, E.R., F.J. MACWILLIAMS & N.J.A. SLOANE, *Gleason's theorem on self-dual codes*, IEEE Trans. Information Theory, IT-18 (1972) 409-414.

[4] BERLEKAMP, E.R. & L.R. WELCH, *Weight distributions of the cosets of the (32,6) Reed-Muller code*, IEEE Trans. Information Theory, IT-18 (1972) 203-207.

[5] BURNSIDE, W., *Theory of groups of finite order*, 2nd edition, 1911; reprinted by Dover, New York, 1955.

[6] CONWAY, J.H., *A group of order 8, 315, 553, 613, 086, 720, 000*, Bull. London Math. Soc., 1 (1969) 79-88.

[7] CONWAY, J.H., *A characterization of Leech's lattice*, Invent. Math., 7 (1969) 137-142.

[8] CONWAY, J.H., *Groups, lattices and quadratic forms*, in: *Computers in algebra and number theory*, SIAM-AMS Proceedings IV, Amer. Math. Soc., Providence, R.I., 1971, pp.135-139.

[9] DYM, H. & H.P. MCKEAN, *Fourier series and integrals*, Acad. Press, New York, 1972.

[10] FEIT, W., *A self-dual even (96,48,16) code*, IEEE Trans. Information Theory, IT-20 (1974) 136-138.

[11] GLEASON, A.M., *Weight polynomials of self-dual codes and the MacWilliams identities*, in: Actes Congrès Internat. Math. 1970, vol. 3, Gauthiers-Villars, Paris, 1971, pp. 211-215.

[12] HOBBS, C.F., *Approximating the performance of a binary group code*, IEEE Trans. Information Theory, IT-11 (1965) 142-144.

[13] JACOBSON, N., *Lectures in abstract algebra, vol. 3*, Van Nostrand, Princeton, N.J., 1964.

[14] KARLIN, M., *New binary coding results by circulants*, IEEE Trans. Information Theory, IT-15 (1969) 81-92.

[15] KASAMI, T. & N. TOKURA, *On the weight structure of Reed-Muller codes*, IEEE Trans. Information Theory, IT-16 (1970) 752-759.

[16] KLEIN, F., *Lectures on the icosahedron and the solution of equations of the fifth degree*, 2nd revised edition, 1913; reprinted by Dover, New York, 1956.

[17] LEECH, J. & N.J.A. SLOANE, *Sphere packings and error-correcting codes*, Canad. J. Math., 23 (1971) 718-745.

[18] LEONT'EV, V.K., *Spectra of linear codes*, in: Third International Symp. on Information Theory, Tallinn, Estonia, June 1973, Abstracts of papers, part II, pp. 102-106.

[19] LINT, J.H. VAN, *Coding theory*, Lecture Notes in Mathematics 201, Springer-Verlag, Berlin, 1971.

[20] MACWILLIAMS, F.J., *Combinatorial problems of elementary abelian groups*, thesis, Dept. of Math., Harvard University, May 1962.

[21] MACWILLIAMS, F.J., *A theorem on the distribution of weights in a systematic code*, Bell System Tech. J., 42 (1963) 79-84.

[22] MACWILLIAMS, F.J. & N.J.A. SLOANE, *Combinatorial coding theory*, to appear.

[23] MACWILLIAMS, F.J., C.L. MALLOWS & N.J.A. SLOANE, *Generalizations of Gleason's theorem on weight enumerators of self-dual codes*, IEEE Trans. Information Theory, IT-18 (1972) 794-805.

[24] MACWILLIAMS, F.J., N.J.A. SLOANE & J.M. GOETHALS, *The MacWilliams identities for nonlinear codes*, Bell System Tech. J., 51 (1972) 803-819.

[25] MACWILLIAMS, F.J., N.J.A. SLOANE & J.G. THOMPSON, *Good self-dual codes exist*, Discrete Math., 3 (1972) 153-162.

[26] MACWILLIAMS, F.J., N.J.A. SLOANE & J.G. THOMPSON, *On the existence of a projective plane of order 10*, J. Combinatorial Theory A, 14 (1973) 66-78.

[27] MALLOWS, C.L. & N.J.A. SLOANE, *An upper bound for self-dual codes*, Information and Control, 22 (1973) 188-200.

[28] MALLOWS, C.L. & N.J.A. SLOANE, *On the invariants of a linear group of order 336*, Proc. Cambridge Philos. Soc., 74 (1973) 435-440.

[29] MALLOWS, C.L. & N.J.A. SLOANE, *Weight enumerators of self-orthogonal codes*, Discrete Math., to appear

[30] MALLOWS, C.L. & N.J.A. SLOANE, unpublished.

[31] MALLOWS, C.L., A.M. ODLYZKO & N.J.A. SLOANE, *Upper bounds for modular forms, lattices and codes*, J. Algebra, to appear.

[32] MILLER, G.A., H.F. BLICHFELDT & L.E. DICKSON, *Theory and applications of finite groups*, 1916; reprinted by Dover, New York, 1961.

[33] MOLIEN, T., *Über die Invarianten der lineare Substitutionsgruppe*, Sitzungsber. König. Preuss. Akad. Wiss., 1897, pp.1152-1156.

[34] PETERSON, W.W. & E.J. WELDON Jr., *Error-correcting codes*, 2nd edition, MIT Press, Cambridge, Mass., 1972.

[35] PLESS, V., *On a new family of symmetry codes and related new five-designs*, Bull. Amer. Math. Soc., 75 (1969) 1339-1342.

[36] PLESS, V., *Symmetry codes over GF(3) and new five-designs*, J. Combinatorial Theory, 12 (1972) 119-142.

[37] PLESS, V., *A classification of self-orthogonal codes over GF(2)*, Discrete Math., 3 (1972) 209-246.

[38] PLESS, V. & J.N. PIERCE, *Self-dual codes over GF(q) satisfy a modified Varshamov bound*, Information and Control, 23 (1973) 35-40.

[39] SHEPHARD, G.C. & J.A. TODD, *Finite unitary reflection groups*, Canad. J. Math., 6 (1954) 274-304.

[40] SLOANE, N.J.A., *Is there a (72,36) d = 16 self-dual code?*, IEEE Trans. Information Theory, IT-19 (1973) 251.

[41] SLOANE, N.J.A. & E.R. BERLEKAMP, *Weight enumerators for second order Reed-Muller codes*, IEEE Trans. Information Theory, IT-16 (1970) 745-751.

[42] SLOANE, N.J.A. & R.J. DICK, *On the enumeration of cosets of first order Reed-Muller codes*, IEEE Internat. Conf. on Commun., Montreal 1971, 7: pp.36-2 to 36-6.

[43] SLOANE, N.J.A. & D.S. WHITEHEAD, *New family of single-error correcting codes*, IEEE Trans. Information Theory, IT-16 (1970) 717-719.

[44] WEYL, H., *The classical groups*, Princeton University Press, Princeton, N.J., 1946.

MATHEMATICAL CENTRE TRACTS 55, 1974, 143-161

THE ASSOCIATION SCHEMES OF CODING THEORY

P. DELSARTE *)

MBLE Research Laboratory, Brussels, Belgium

1. INTRODUCTION

This paper contains the bases of an algebraic theory of certain association schemes, called *polynomial schemes*. Special emphasis is put on concepts arising from the theories of error correcting codes and of combinatorial designs. The main goal is to provide a general framework in which various applications can be treated by similar methods. In this respect, an interesting formal duality is exhibited between non-constructive coding and design theory.

First, in section 2, some general definitions and preliminary results are given about an *association scheme* (for short, a scheme), especially from the point of view of its *Bose-Mesner algebra* [3]. The natural schemes of coding theory are defined and some polynomial properties of their Bose-Mesner algebras are emphasized.

Section 3 is devoted to the concept of *inner* and *outer distribution* of a subset in an association scheme. Essentially, it is shown that the inner distribution of any subset satisfies certain well-defined inequalities. This result leads to *linear programming problems* having interesting applications in classical theory of codes and designs. Useful relations between the inner and outer distributions are also obtained.

In section 4, we give an axiomatic definition of *polynomial schemes*, which generalize the "coding schemes". In this context, several results about generalized codes and designs are presented. To be more specific, let us mention the questions of *perfect codes* and *tight designs*, among others.

Finally, in section 5, as an application, we briefly consider the *linear codes*, for which certain aspects of the general theory have simple interpretations.

The present paper essentially constitutes part of the author's recent

---

*) The author's participation in this meeting was not supported by NATO.

work [8], where proofs of all theorems given below can be found. A few proofs are incorporated herein, mainly to illustrate the methods.

2. DEFINITIONS AND PRELIMINARIES

Before examining more general notions, we shall briefly describe a suitable framework for classical coding theory. Let F be a finite alphabet, of cardinality $q \geq 2$. Then, for a given $n \geq 1$, the set $X = F^n$ of all n-tuples over F is made a metric space (X,d_H) by definition of the *Hamming distance*:

$$d_H(x,y) = |\{\nu \mid 1 \leq \nu \leq n,\ x_\nu \neq y_\nu\}|, \quad x = (x_\nu)_{\nu=1}^n,\ y = (y_\nu)_{\nu=1}^n.$$

(The distance between two n-tuples is the number of coordinate positions in which they differ.) A q-*ary code of length* n then simply is any non-empty subset of X, endowed with the Hamming metric.

Let us now define the set $R = \{R_0,R_1,\ldots,R_n\}$ of distance relations on X:

$$R_i = \{(x,y) \in X^2 \mid d_H(x,y) = i\}, \quad i=0,1,\ldots,n.$$

We shall call the pair (X,R) a *Hamming scheme*, using the notation H(n,q) for it. (Another terminology is "hypercubic type association schemes"; we refer to YAMAMOTO, FUJII & HAMADA [28].) The following properties of H(n,q) are easily checked:

A1. *The set* R *is a partition of* X^2, *the* R_i *are symmetric (i.e.* $R_i^{-1} = R_i$ *for all* i) *and* R_0 *is the diagonal (=* $\{(x,x) \mid x \in X\}$).

A2. *Let* $(x,y) \in R_k$. *The integer* $p_{i,j,k} = |\{z \in X \mid (x,z) \in R_i,\ (y,z) \in R_j\}|$ *is a constant; it only depends on* (i,j,k).

We now turn to a more general situation. Let X be a finite set of cardinality ≥2, and, for a given integer $n \geq 1$, let $R = \{R_0,R_1,\ldots,R_n\}$ be a set of n+1 relations R_i on X satisfying the axioms A1 and A2. Then (X,R) is called an *association scheme with* n *classes*. This is the concept introduced by BOSE & SHIMAMOTO [4]. (According to HIGMAN's terminology, it is a homogeneous coherent configuration with trivial pairing [13].) Two points $x,y \in X$ are said to be i-*th associates* whenever $(x,y) \in R_i$ holds. The $p_{i,j,k}$ are the *intersection numbers* of the scheme. They satisfy the symmetry relations $p_{i,j,k} = p_{j,i,k}$; for less trivial identities, the reader is referred to BOSE & MESNER [3].

The case n = 2 corresponds to the *strongly regular graphs* introduced by BOSE [2]. In general, an association scheme (X,R) may be viewed as a complete graph on the vertex set X, with n distinct colours $c_1,\ldots,c_n$ for the edges: the edge (x,y) is coloured in c_i if and only if x and y are i-th associates, for i=1,2,...,n.

REMARKS

(i) Let G be a transitive permutation group on a set X, with $2 \le |X| < \infty$, and let $R = \{R_0,R_1,\ldots,R_n\}$ denote the set of all G-orbits of X^2. Then, provided the R_i are symmetric, it is well-known that (X,R) is an association scheme. This is called the *group case* (cf. HIGMAN [13]). For instance, the Hamming schemes belong to the group case.

(ii) Assume that, for a given scheme (X,R), two points $x,y \in X$ are i-th associates if and only if they are at distance $\rho(x,y) = i$ in the graph (X,R_1). Then (X,R_1) is called a *perfectly regular graph* (cf. HIGMAN [13]) or a *metrically regular graph* (cf. DOOB [9]). It can easily be seen that an association scheme has such a property of being "generated by a graph" if and only if the intersection numbers satisfy $p_{i,j,k} \neq 0$ whenever k = i+j and

$$(p_{i,j,k} \neq 0) \Rightarrow (|i-j| \le k \le i+j).$$

In this case, we call (X,R) a *metric scheme*. The Hamming schemes clearly are metric, with $\rho = d_H$.

Let R_i be any relation on X. We shall denote by D_i the *adjacency matrix* of (X,R_i), i.e. the square matrix of order $|X|$ over $\mathbb{R}$, having X as row and column labeling set, whose (x,y)-entry is

$$D_i(x,y) = \begin{cases} 1 & \text{if } (x,y) \in R_i, \\ 0 & \text{otherwise,} \end{cases}$$

for all $x,y \in X$. The axioms A1, A2 for an association scheme (X,R) can now be expressed in matrix form as follows:

A'1. $\sum_i D_i = J$ (= *all-one matrix*), $D_i^T = D_i$ *for all* i, $D_0 = I$.

A'2. $D_iD_j = D_jD_i = \sum_k p_{i,j,k}D_k$ *for all* i,j.

As an immediate consequence, we have the following result, due to BOSE & MESNER [3]: Let R be a set of n+1 relations R_i on X, satisfying A1. Then (X,R) is an association scheme, with n classes, if and only if the adjacency matrices D_i generate an (n+1)-dimensional commutative algebra over $\mathbb{R}$.

We shall now examine in some detail the properties of the algebra $A = \langle D_0, D_1, \ldots, D_n \rangle$ of an association scheme (X,R), which we call the *Bose-Mesner algebra*. It is easy to show that A is a semisimple algebra, isomorphic to $\mathbb{R}^{n+1}$. In other words, A admits a unique basis $(J_0, J_1, \ldots, J_n)$ of *minimal idempotents* J_k, being mutually orthogonal: $J_i J_k = \delta_{i,k} J_k$, and satisfying $\sum J_k = I$. Moreover, for a suitable numbering, we have $J_0 = |X|^{-1} J$. This structure has been described first by OGAWA [22].

Let us write the expansion of the adjacency matrices D_k with respect to the basis of minimal idempotents J_i:

$$D_k = \sum_{i=0}^{n} P_k(i) J_i, \qquad k=0,1,\ldots,n, \tag{2.1}$$

for some real numbers $P_k(i)$, uniquely defined. These are the eigenvalues of D_k. Indeed (2.1) yields $D_k J_i = P_k(i) J_i$. This shows, in addition, that the column spaces of the J_i are the common eigenspaces of all matrices belonging to A. We shall denote by μ_i, and call i-th *multiplicity*, the dimension of the i-th eigenspace, i.e. $\mu_i = \operatorname{rank}(J_i)$, i=0,1,...,n. We also point out that the eigenvalue $v_k = P_k(0)$ has an obvious combinatorial meaning, namely v_k = *valency* of R_k = number of k-th associates of a fixed point in X (= number of ones in each row of D_k).

It is often more interesting to characterize an association scheme by the eigenvalues $P_k(i)$ rather than by the intersection numbers $p_{i,j,k}$, although either set of parameters can be derived from the other. In fact we shall make use of the *eigenmatrices* P and Q: these are the non-singular square matrices of order n+1 defined to be

$$P = [P_k(i) : i,k \in \{0,1,\ldots,n\}], \qquad Q = |X| P^{-1}.$$

So $P_k(i)$ is the (i,k)-entry of P. Similarly, $Q_k(i)$ will denote the (i,k)-entry of Q, for i,k=0,1,...,n. Thus, according to (2.1), we have

$$|X| J_k = \sum_{i=0}^{n} Q_k(i) D_i, \qquad k=0,1,\ldots,n. \tag{2.2}$$

It is easily seen that $Q_0(i) = 1$ and $Q_k(0) = \mu_k$ hold, for all i,k.

Given an (n+1)-tuple $\underline{c} = (c_0, c_1, \ldots, c_n)$ of real numbers c_i, we shall

denote by $\Delta_{\underline{c}}$ the diagonal matrix $\mathrm{diag}(c_0,c_1,\ldots,c_n)$. The following theorem, due to YAMAMOTO, FUJII & HAMADA [28], exhibits interesting *orthogonality relations* on the eigenmatrices.

THEOREM 1. $\quad Q^T\Delta_{\underline{v}}Q = |X|\Delta_{\underline{\mu}}, \quad P^T\Delta_{\underline{\mu}}P = |X|\Delta_{\underline{v}}.$

PROOF. One easily obtains the first identity by expressing $J_iJ_k = \delta_{i,k}J_k$ in the basis of matrices D_j. The details will not be given. Then the second identity follows from the first one, by use of $P = |X|Q^{-1}$. □

When the eigenmatrices P and Q are interchanged, the relations of theorem 1 are transformed into each other. This is the first occurrence of a nice *formal duality* appearing at several places in the theory. In this respect, one may ask the question whether there exists a *dual scheme* of (X,R), that is, an association scheme (X',R') whose eigenmatrices are P' = Q and Q' = P. Such an "actual duality" can indeed by defined for certain association schemes, such as those admitting an Abelian regular automorphism group. (In this case the Bose-Mesner algebras are in fact *Schur rings*. The duality introduced by TAMASCHKE [25] for Schur rings can be used to define dual association schemes.) For instance, the Hamming schemes admit a dual. More precisely, they are self-dual so that one has P = Q, ie. $P^2 = q^nI$.

For the Hamming schemes, it turns out that the eigenvalue $P_k(i)$ can be expressed as a polynomial of degree k with respect to i:

$$P_k(i) = \sum_{j=0}^{k} (-q)^j(q-1)^{k-j}\binom{n-j}{k-j}\binom{i}{j}, \quad k=0,1,\ldots,n.$$

In fact, the $P_k(z)$ constitute a well-known class of orthogonal polynomials, first introduced by KRAWTCHOUCK [16]. Theorem 1, with $v_k = \mu_k = \binom{n}{k}(q-1)^k$, precisely contains the orthogonality relations on the *Krawtchouck polynomials*:

$$\sum_{k=0}^{n} P_i(k)P_j(k)\binom{n}{k}(q-1)^k = q^n\binom{n}{i}(q-1)^i\delta_{i,j}.$$

It can be shown that the association schemes whose eigenmatrix P has such "polynomial properties" are exactly the metric schemes. They will be studied in section 4.1.

We conclude the present section with a definition of a class of metric schemes that constitute a natural framework for the theory of *constant weight binary codes*, first considered by JOHNSON [15]: for F = {0,1} and an integer $v \geq 2$, the weight of an element $x \in F^v$ is the number of coordinates x_ν being

equal to 1. Let X denote the set of elements of a given weight n in F^v, with $1 \le n \le v/2$. For k=0,1,...,n, we define the distance relations

$$R_k = \{(x,y) \in X \mid d_H(x,y) = 2k\}.$$

Then it is easily verified that (X,R) is an association scheme, with n classes, for $R = \{R_0,\ldots,R_n\}$. We shall call it a *Johnson scheme*, using the notation J(n,v). (The classical terminology is "triangular type association scheme".)

It turns out that, for the Johnson scheme J(n,v), the eigenvalue $P_k(i)$ can be expressed as a polynomial of degree k with respect to the variable $z_i = i(v+1-i)$. Explicit formulae for $P_k(i)$ have been discovered by OGASAWARA [21] and by YAMAMOTO, FUJII & HAMADA [28]:

$$P_k(i) = \sum_{j=0}^{k} (-1)^{k-j} \binom{n-j}{k-j} \binom{n-i}{j} \binom{v-n+j-i}{j}.$$

On the other hand, it can be shown that the elements of the second eigenmatrix have similar properties: $Q_k(i)$ is a polynomial of degree k in the variable $z_i = i$, for all k. Notice that, according to theorem 1, the P- and Q-polynomials of the scheme J(n,v) form two families of orthogonal polynomials, with $v_k = \binom{n}{k}\binom{v-n}{k}$ and $\mu_k = \binom{v}{k} - \binom{v}{k-1}$.

3. DISTRIBUTION OF A SUBSET IN AN ASSOCIATION SCHEME

Let Y be a non-empty subset of X for any given association scheme (X,R). We define the *inner distribution* of Y to be the (n+1)-tuple $\underline{a} = (a_0,a_1,\ldots,a_n)$ of rational numbers a_i given by

$$|Y|a_i = |R_i \cap Y^2|, \quad i=0,1,\ldots,n.$$

Thus, a_i is the average number of points of Y being i-th associates of a fixed point of Y. Clearly, $a_0 = 1$ and $\sum a_i = |Y|$ hold. For a metric scheme, Y is called a *code* in (X,R) and $\underline{a}$ is the *distance distribution* of the code.

A central question in the rest of this paper will be the following. What can be said about a subset (a "code", a "design") when its inner distribution is given? First, however, we wish to characterize those (n+1)-tuples that are inner distributions of subsets. We shall now establish a very useful result in this direction.

THEOREM 2. *The inner distribution* $\underline{a}$ *of any subset* $Y \subseteq X$ *satisfies* $\underline{a}Q_k \geq 0$, *for all* k, *where* $Q = [Q_0, Q_1, \ldots, Q_n]$ *is the second eigenmatrix of the association scheme* (X,R).

PROOF. We shall make use of the vector $\underline{u} \in \mathbb{R}(X)$, characterizing Y as a subset of X, defined by $u(x) = 1$ or 0 according to whether $x \in Y$ or $x \in X-Y$. Then we clearly have $|Y|a_i = \underline{u}^T D_i \underline{u}$. Hence, using (2.2), we deduce

$$|Y|\underline{a}Q_k = \underline{u}^T\Big(\sum_{i=0}^{n} Q_k(i)D_i\Big)\underline{u} = |X|\underline{u}^T J_k \underline{u}.$$

Now the idempotent matrix J_k is positive semi-definite. Therefore, the right member is ≥ 0 and the theorem is proved. □

REMARK. Let us briefly indicate an interpretation of theorem 2 in the theory of linear codes. When the alphabet F is a field $GF(q)$, the set $X = F^n$ has the structure of a vector space over F. Then a q-ary code of length n is called *linear* whenever it is a subspace of X. The distance distribution $\underline{a}$ of a linear code with respect to the Hamming scheme simply is the classical *weight distribution*: a_i is the number of codevectors having weight i. (The Hamming weight of $x \in X$ is the number of non-zero components x_ν.) The *dual* Y' of a linear code Y is the set of vectors $x \in X$ such that $x_1y_1 + \ldots + x_ny_n = 0$ holds for every $y \in Y$. Clearly, Y' is itself linear, with $\dim(Y') = n - \dim(Y)$. Then the weight distributions $\underline{a}$ and $\underline{a}'$ of Y and Y' are related by $|Y|\underline{a}' = \underline{a}Q$, where $Q = P$ is the matrix of Krawtchouck polynomials $P_k(z)$. This is a version of the celebrated *MacWilliams identities* [20] on the weight distributions of dual linear codes. Thus, in the linear case, theorem 2 reduces to the trivial property $a_k' \geq 0$. We have seen that this remains valid for unrestricted codes in Hamming schemes when $\underline{a}'$ is defined as the formal "Krawtchouck-MacWilliams transform" of the distance distribution $\underline{a}$.

Theorem 2 leads to *linear programming problems* in the theory of subsets $Y \subseteq X$ whose specific properties can be expressed in terms of linear relations on the inner distribution with respect to a given association scheme. One is interested in upper or lower bounds on the cardinality of subsets satisfying these conditions. The linear programming problems in the $(n+1)$-tuple $\underline{a} = (a_0, a_1, \ldots, a_n)$ of real variables a_i, have the following form:

$$\begin{cases} \sum_{i=0}^{n} f_{i,j} a_i = 0, \quad j=1,2,\ldots,m, \\ a_0 = 1, \quad a_i \geq 0, \quad \underline{a}Q_k \geq 0, \quad \forall\, i,k, \\ \text{maximize (or minimize) } g = \sum_{i=0}^{n} a_i . \end{cases}$$

The first line contains the specifications of the problem, whereas the second line contains general necessary conditions (cf. theorem 2). So, since $g = |Y|$ holds, each subset Y under consideration satisfies the *linear programming bounds*

$$\min(g) \leq |Y| \leq \max(g).$$

In fact, we are mainly interested in two types of applications (i.e. in two types of matrices $[f_{i,j}]$), that are "dual to each other":

(i) *The codes with specified distance* δ, characterized by $a_1 = a_2 = \ldots = a_{\delta-1} = 0$, for a given $\delta \geq 1$.

(ii) *The designs with specified strength* τ, characterized by $\underline{a}Q_1 = \underline{a}Q_2 = \ldots = \underline{a}Q_\tau = 0$, for a given $\tau \geq 0$.

The significance of problem (i) of *δ-codes* in metric schemes is obvious and needs no comment. As for the problem (ii) of *τ-designs* (cf. section 4.2), we can give no general "combinatorial interpretation" of it. However, for the Hamming and Johnson schemes, we have the following result.

THEOREM 3. *A τ-design Y in H(n,q) is equivalent to an orthogonal array $[\lambda q^\tau, n, q, \tau]$, without repeated columns, of strength τ, having n constraints, index λ, in q symbols, with $|Y| = \lambda q^\tau$.* (This is the concept introduced by RAO [23].) *A τ-design Y in J(n,v) is equivalent to a classical τ-design $S_\lambda(\tau,n,v)$, without repeated blocks, on v points, block size n, with $|Y| = \lambda \binom{v}{\tau} / \binom{n}{\tau}$.* (This is the concept introduced by HANANI [12] and HUGHES [14].)

PROOF. We shall only consider the case of Johnson schemes: $(X,R) = J(n,v)$. Let X_i denote the subset of $\{0,1\}^v$ formed by all elements of weight i, for $i=0,1,\ldots,n$. Next, let Y be a non-empty subset of X $(= X_n)$. Given $z \in X_i$, we shall denote by $\lambda_i(z)$ the number of "blocks" $y \in Y$ such that $d_H(y,z) = n-i$. Then the average value of $\lambda_i(z)$ over X_i is equal to

$$\lambda_i = |Y| \binom{n}{i} / \binom{v}{i}, \quad i=0,1,\ldots,n.$$

By definition, Y forms a τ-design $S_\lambda(\tau,n,v)$, with $\lambda = \lambda_\tau$, if and only if $\lambda_i(z)$ is a constant $(= \lambda_i)$, for all $z \in X_i$, whenever $i=0,1,\ldots,\tau$.

On the other hand, it is easy to show, by a counting argument, that

$$\sum_{z\in X_i} (\lambda_i(z) - \lambda_i)^2 = |Y|\left\{ \sum_{j=0}^{n} a_j \binom{n-j}{i} - \lambda_i \binom{n}{i} \right\}$$

holds, for all $i \le n$, where $\underline{a}$ is the distance distribution of Y. Therefore, Y forms a τ-design if and only if the right member vanishes, i.e.

$$\sum_{j=0}^{n} a_j \binom{n-j}{i} = \lambda_i \binom{n}{i},$$

for $i=0,1,\ldots,\tau$. Observing that X itself forms a trivial τ-design, and that the distance distribution $\underline{v}$ of X is given by the valencies v_j, we can also write this as follows

$$\sum_{j=0}^{n} \{|X|a_j - |Y|v_j\}\binom{n-j}{i} = 0.$$

Since the polynomials $\binom{n-z}{i} \in \mathbb{R}[z]$, with $i=0,1,\ldots,\tau$, form a basis of the vector space of polynomials of degree $\le\tau$, the above system is equivalent to

$$\sum_{j=0}^{n} \{|X|a_j - |Y|v_j\}Q_k(j) = 0,$$

for $k=0,1,\ldots,\tau$. (We have used the property $\deg(Q_k(z)) = k$ of the eigenmatrix Q.) From the orthogonality relations on Q we finally obtain the desired characterization of a τ-design, namely $\sum_j a_j Q_k(j) = |Y|\delta_{0,k}$ for all $k \le \tau$. □

According to theorem 3, we may use the linear programming method in order to obtain a lower bound on the index λ of orthogonal arrays with given τ,n,q and of ordinary τ-designs with given τ,n,v. It can be shown that the linear programming bound improves the RAO inequality [23] for orthogonal arrays as well as the FISHER-PETRENJUK-WILSON inequality [27] for τ-designs (cf. section 4.2).

Similarly, the linear programming method implies several classical bounds for the problem of δ-codes in the Hamming and Johnson schemes, such as the elementary sphere packing bound, the Singleton bound and the Plotkin bound (for references, cf. [8]). A major interest of the method lies in the fact that, by use of duality in linear programming, it yields strong characterizations for codes achieving the bounds. In particular, the Lloyd theorem for perfect codes can be obtained in this manner (see section 4.1).

Let us finally introduce the *outer distribution* of a non-empty subset $Y \subseteq X$ with respect to a given association scheme (X,R): it is the integer matrix B, having X and $\{0,1,\ldots,n\}$ as row and column labeling sets, respectively, the (x,i)-entry being defined as

$$B(x,i) = |R_i \cap (\{x\} \times Y)|,$$

for $x \in X$ and $i=0,1,\ldots,n$. Thus, $B(x,i)$ is the number of points in Y that are i-th associates of the fixed point x. We shall now establish a useful relation between the inner distribution $\underline{a}$ and the outer distribution B of any $Y \subseteq X$. Like before, P and Q denote the eigenmatrices and $\Delta_{\underline{c}}$ stands for $\mathrm{diag}(c_0,c_1,\ldots,c_n)$.

THEOREM 4. $B^TB = |X|^{-1}|Y|P^T\Delta_{\underline{a}Q}P$.

PROOF. Counting in two different ways the number of triples $(x,y,y') \in X \times Y \times Y$ such that $(x,y) \in R_i$ and $(x,y') \in R_j$ we obtain the identity

$$(B^TB)(i,j) = |Y| \sum_{k=0}^{n} p_{i,j,k}a_k.$$

Define $\underline{b} = \underline{a}Q$, whence $|X|\underline{a} = \underline{b}P$. Using the formulas $P_i(u)P_j(u) = \sum_k p_{i,j,k}P_k(u)$ deduced from axiom A'2 we readily deduce

$$(B^TB)(i,j) = |X|^{-1}|Y| \sum_{u=0}^{n} b_uP_i(u)P_j(u),$$

which is the desired result. □

It follows from theorem 4 that the rank of the matrix B is equal to the number of non-zero components of the vector $\underline{a}Q$. Notice also that, since B^TB is positive semi-definite, we have obtained a new proof of theorem 2: all components of $\underline{a}Q$ are non-negative.

Let $Q_0,Q_1,\ldots,Q_n$ be the columns of Q. Then theorem 4 clearly implies the identity

$$\|BQ_k\|^2 = |X||Y|\underline{a}Q_k, \quad 0 \le k \le n,$$

where $\|\ .\ \|$ denotes the euclidean norm. This not only shows that $\underline{a}Q_k \ge 0$ holds, but also that the vector BQ_k is zero if and only if $\underline{a}Q_k$ is zero. This property is very useful for actual computation of the outer distribution B. In certain cases, it allows to determine B from the inner

distribution a. More details about this matter are given below, in the context of metric schemes. The reader shall have noticed the significance of the outer distribution in the classical theory of linear codes: the rows of B are the weight distributions of the cosets of the given code (cf. section 5).

4. POLYNOMIAL SCHEMES

We have mentioned the polynomial properties of the eigenmatrices P and Q for the Hamming and Johnson schemes occurring in the classical theory of codes and designs. In the present section, we shall take these properties as axioms and set up the bases of a general theory of "codes and designs in polynomial schemes". The main idea consists in trying to derive as much information as possible about a code (or a design) from its inner distribution, and, more precisely, from certain fundamental parameters depending on the inner distribution.

DEFINITION. Let $z_0, z_1, \ldots z_n$ be distinct non-negative real numbers, with $z_0 = 0$. Assume that the entries of the eigenmatrix P can be written as

$$P_k(i) = \Phi_k(z_i), \quad i,k=0,1,\ldots,n,$$

where $\Phi_k(z) \in \mathbb{R}[z]$ is a polynomial of degree k. Then the given association scheme is said to be P-*polynomial* with respect to the z_i. A Q-polynomial scheme is defined analogously from the properties of the eigenmatrix Q.

We recall that the Hamming and Johnson schemes are P-polynomial for $z_i = i$ and $z_i = i(v+1-i)$, respectively. They both are Q-polynomial for $z_i = i$. In fact, there exist several other families of association schemes having the P- and Q-polynomial properties. Some of them also have interesting applications in coding theory.

The orthogonality relations of theorem 1 can be interpreted as follows. For a P-polynomial scheme, $\Phi_0(z), \Phi_1(z), \ldots, \Phi_n(z)$ form a class of orthogonal polynomials over the set $\{z_0, z_1, \ldots, z_n\}$, the weight function w being given by $w(z_i) = \mu_i$. Moreover, it can be shown that the *sum polynomials*

$$\Psi_k(z) = \Phi_0(z) + \Phi_1(z) + \ldots + \Phi_k(z),$$

with $k=0,1,\ldots,n-1$, form a class of orthogonal polynomials over the set

$\{z_1,z_2,\ldots,z_n\}$ for the weight function $w'(z) = zw(z)$. The dual results hold in the theory of Q-polynomial schemes, the weights being $w(z_i) = v_i$ and $w'(z_i) = z_i v_i$.

4.1. P-polynomial schemes

It turns out that a given association scheme is P-polynomial if and only if it is metric, as defined in section 2. Thus the mapping ρ, of X^2 into $\{0,1,\ldots,n\}$, given by $\rho(x,y) = i$ whenever $(x,y) \in R_i$, is a *distance function* on X. (It is the natural distance in the graph (X,R_1).)

A *code* Y, that is, a non-empty subset $Y \subseteq X$ in a metric scheme (X,R), will be characterized by two parameters, d and r, both deduced from the distance distribution $\underline{a}$:

(i) The *minimum distance* d is the largest integer ≥ 1 such that $a_i = \delta_{0,i}$ for $i=0,1,\ldots,d-1$.

(ii) The *external distance* r is the number of non-zero components of the n-tuple $(\underline{a}Q_1,\underline{a}Q_2,\ldots,\underline{a}Q_n)$.

We shall only consider *non-trivial codes*, i.e. proper subsets of X containing at least two elements. Then the parameters satisfy $1 \leq r,\ d \leq n$. The significance of d is obvious: it is the smallest positive value assumed by the distance function $\rho(x,y)$ for $x,y \in Y$. As for the external distance r, our terminology is based on the following result.

THEOREM 5. *Each point of* X *is at distance* $\leq r$ *from at least one point belonging to the code* Y:

$$\min_{y\in Y} \rho(x,y) \leq r, \quad \forall x \in X.$$

PROOF. Let us first define the following two subsets of $\{0,1,\ldots,n\}$, with equal cardinalities: $K = \{0,1,\ldots,r\}$ and $L = \{k \mid \underline{a}Q_k \neq O\}$. We shall denote by $\bar{B}$, $\bar{P}$ and $\bar{\Delta}$ the restrictions of the matrices B, P and $\Delta_{\underline{a}Q}$ to the sets $X \times K$, $L \times K$ and $L \times L$, respectively. Then the equation of theorem 4 implies

$$\bar{B}^T\bar{B} = |X|^{-1}|Y|\bar{P}^T\bar{\Delta}\bar{P}.$$

By definition, $\bar{\Delta}$ is a non-singular diagonal matrix. On the other hand, from the P-polynomial property it clearly follows that $\bar{P}$ also is non-singular. Hence $\bar{B}^T\bar{B}$ is non-singular and we deduce $\operatorname{rank}(\bar{B}) = \operatorname{rank}(B) = r+1$. In other

words, the columns $B_0, B_1, \ldots, B_r$ of $\bar{B}$ form a basis for the column space of B. So, since a row B(x) of B cannot be identically zero, at least one of the integers $B_0(x), B_1(x), \ldots, B_r(x)$ must be non-zero, for every given $x \in X$. Remembering that $B_i(x)$ is the number of points $y \in Y$ such that $\rho(x,y) = i$, we obtain the desired result. □

It must be observed that, in general, r is not the "true external distance": there may exist no point $x \in X$ at minimum distance r from the code Y. As an easy consequence of theorem 5 we deduce an interesting inequality, first discovered by MACWILLIAMS [19] for linear codes.

THEOREM 6. *For any code,* $\lfloor (d-1)/2 \rfloor \leq r$ *holds.*

PROOF. Let us define the spheres $S_e(y) = \{x \in X \mid 0 \leq \rho(x,y) \leq e\}$, of radius $e = \lfloor (d-1)/2 \rfloor$, centred at the points $y \in Y$. By definition, these spheres are mutually disjoint. So, any point $x_0 \in X$ at distance e from some $y \in Y$ is at minimum distance e from Y. According to theorem 5, this implies $e \leq r$. □

Moreover, it can be shown that the equality e = r holds if and only if the spheres $S_e(y)$ form a partition of X, for y running through Y. In this case, by extension of the classical notion, Y is called a *perfect code of order* e. This concept has also been independently introduced and investigated by BIGGS [1] in the context of distance transitive graphs (which corresponds to the group case for metric schemes).

In connection with theorem 6, let us mention the following two bounds on the cardinality of any code with given parameters d and r:

$$(4.1) \qquad \sum_{i=0}^{e} v_i \leq |Y|^{-1}|X| \leq \sum_{i=0}^{r} v_i,$$

with $e = \lfloor (d-1)/2 \rfloor$. The left member is the obvious sphere packing bound. The right member is the "covering bound", which easily follows from theorem 5. It turns out that, if one of the bounds is achieved, then so is the other. Therefore, we may take either equality in (4.1) as a definition for perfect codes.

Let us now mention, without proof, a generalized version of the *Lloyd theorem* [1,7,8,17] which yields a strong necessary condition for perfect codes.

THEOREM 7. *Let there exist a perfect code of order* e *in a* P*-polynomial (= metric) scheme. Then the sum polynomial* $\Psi_e(z)$ *admits* e *distinct zeros in the set* $\{z_1, z_2, \ldots, z_n\}$, *namely those* z_k *such that* $\underline{a}Q_k \neq 0$.

From the fact that $\Psi_0(z), \Psi_1(z), \ldots, \Psi_{n-1}(z)$ form a class of orthogonal polynomials, it is possible to derive explicit expressions for the distance distribution of a perfect code. In fact, it turns out that the full outer distribution of perfect codes of a given order only depends on the parameters of the scheme.

Unfortunately, there are "very few" perfect codes in the classical coding schemes. (We refer to VAN LINT [18].) Let us now define a weaker property, which however seems interesting: the complete regularity. A code Y is called *completely regular* if the row B(x) of its outer distribution only depends on the minimum distance between x and Y, for all $x \in X$. The following theorem contains a sufficient condition for this property, in terms of the fundamental parameters.

THEOREM 8. *If the minimum distance* d *and the external distance* r *of a code satisfy* $d \geq 2r-1$, *then the code is completely regular.*

This result not only applies to perfect codes ($d = 2r+1$), but also to the nearly perfect codes defined by GOETHALS & SNOVER [11] and to the uniformly packed codes introduced by SEMAKOV, ZINOV'EV & ZAITZEV [24].

4.2. Q-polynomial schemes

The theory is formally similar to that of the preceding section. One would of course like to know an intrinsic interpretation of the concept of Q-polynomial schemes (i.e. the "dual" of the metric property). Unfortunately, this is an unsolved question, although some useful algebraic criteria are known for the Q-polynomial property.

We shall now investigate the dual notion of a code in a metric scheme. A *design* Y is a non-empty subset $Y \subseteq X$ for a Q-polynomial scheme (X,R); it will be characterized by two fundamental parameters, t and s, deduced from the inner distribution $\underline{a}$:

(i) The *maximum strength* t is the largest integer ≥ 0 such that $\underline{a}Q_k = |Y|\delta_{0,k}$ for $k=0,1,\ldots,t$.

(ii) The *degree* s is the number of non-zero components of the n-tuple $(a_1,a_2,\ldots,a_n)$.

We shall only consider non-trivial designs (i.e. assume $1 < |Y| < |X|$). Then $1 \leq s,\ t+1 \leq n$ holds. The interpretation of the degree is clear: s is the number of distinct colours appearing in the subgraph of (X,R) whose vertex set is Y. The meaning of the maximum strength is less obvious; it must be discovered in each particular case. For the Hamming and Johnson schemes, the significance of t-designs of strength t has been emphasized in theorem 3. This motivates a study of t-designs in general Q-polynomial schemes.

The following result is dual to theorem 6 about codes in metric schemes (for the correspondence $t \leftrightarrow d-1$, $s \leftrightarrow r$).

THEOREM 9. *For any design,* $\lfloor t/2 \rfloor \leq s$ *holds.*

PROOF. Let $\underline{a}$ be the inner distribution of Y. Assume $s < e = \lfloor t/2 \rfloor$. Then there exists a polynomial $f(z) \in \mathbb{R}[z]$, of degree e, vanishing at each point z_i such that $a_i \neq 0$, for $i=0,1,\ldots,n$. Consider the expansion of $(f(z))^2$, which has degree $\leq t$, in the basis of polynomials $\Phi_k(z)$ associated to the eigenmatrix Q:

$$(f(z))^2 = \sum_{k=0}^{t} b_k \Phi_k(z).$$

(The real numbers b_k are uniquely derived from the values $f(z_i)$ by the formulas $|X|b_k = \sum_i P_i(k)(f(z_i))^2$.) Using $\Phi_k(z_i) = Q_k(i)$, we obtain

$$\sum_{i=0}^{n} a_i (f(z_i))^2 = \sum_{k=0}^{t} b_k (\underline{a}Q_k) = b_0|Y|.$$

Now, by definition of f(z), the left member is zero, whereas the right member, being equal to $|X|^{-1}|Y| \sum_i v_i (f(z_i))^2$, is strictly positive. The desired inequality $e \leq s$ follows from this contradiction. □

We have seen that $\lfloor (d-1)/2 \rfloor = r$ can be taken for the definition of a perfect code of order e in a metric scheme. Analogously, we take the equality $\lfloor t/2 \rfloor = s$ as a definition of a *tight design* of degree s in a Q-polynomial scheme. It turns out that this coincides with the concept introduced by WILSON [26] for classical t-designs (in the Johnson schemes). In the case of Hamming schemes, a tight design is equivalent to a *generalized Hadamard code* [7].

The result of theorem 9 can also be viewed as a direct consequence of the following two inequalities (dual to those of (4.1)):

$$(4.2) \qquad \sum_{i=0}^{e} \mu_i \le |Y| \le \sum_{i=0}^{s} \mu_i,$$

with $e = \lfloor t/2 \rfloor$. The left bound is due to RAO [23] for the Hamming schemes and to WILSON & RAY-CHAUDHURI [27] for the Johnson schemes (where it takes the simple form $|Y| \ge \binom{v}{e}$). As for the right bound, it has been first discovered by WILSON [26] and by the author [7] in the case of Johnson and Hamming schemes, respectively.

It turns out that, if one of the bounds (4.2) is achieved, then so is the other. Therefore, we may take either equality in (4.2) to define tight designs. Let us also mention a dual of the Lloyd theorem, extending known results on classical tight designs [26] and generalized Hadamard codes [7].

THEOREM 10. *Let there exist a tight design of degree* s *in a* Q*-polynomial scheme. Then the sum polynomial* $\Psi_s(z)$ *admits* s *distinct zeros in the set* $\{z_1, z_2, \ldots, z_n\}$, *namely those* z_k *such that* $a_k \neq 0$.

The above result implies that the s colours of the subgraph of (X,R) having Y as vertex set are determined from the parameters of the scheme, when Y is a tight design. Moreover, it can be shown that this subgraph itself is a Q-polynomial association scheme, with s classes. This is a particular case of a more general result (to be compared with theorem 8):

THEOREM 11. *If the maximum strength* t *and the degree* s *of a design satisfy* $t \ge 2s-2$, *then the design carries a* Q*-polynomial scheme, with* s *classes.*

For s = 2, this theorem has been discovered first by GOETHALS & SEIDEL [10] in the case of Johnson schemes and by the author [6] in the case of Hamming schemes (at least for linear codes). This led to some interesting constructions of strongly regular graphs. Recently, CAMERON [5] also obtained theorem 11 for designs of any degree s in a Johnson scheme.

5. APPLICATION. LINEAR CODES IN HAMMING SCHEMES

In the preceding section we have exhibited a formal duality between the theories of P- and Q-polynomial schemes, between the concepts of codes and designs, both characterized by a pair of fundamental parameters. In

Hamming schemes this duality becomes actual for the class of *linear codes* (or designs). We recall that the eigenmatrices P and Q of a Hamming scheme H(n,q) are equal, and they are determined from the Krawtchouk polynomials.

For a given linear code Y of length n over GF(q), the *Hamming weight* of an n-tuple over F = GF(q) is defined to be the number of its non-zero components. Let $w_1, w_2, \ldots, w_s$ be the values assumed by the Hamming weight over non-zero elements of Y. The w_i are called the *weights* of Y. The degree of Y clearly is equal to the number of distinct weights, and the minimum distance is equal to the minimum weight.

We shall denote by Y' the dual of the linear code Y. Then the respective weight distributions $\underline{a}$ and $\underline{a}'$ of Y and Y' are related by the *MacWilliams identities* $|Y|\underline{a}' = \underline{a}Q$. Let d,r,t,s be the four fundamental parameters of Y and d',r',t',s' the corresponding parameters of Y'. Then it immediately follows from the MacWilliams identities that we have

$$d' = t+1 \quad , \quad r' = s \quad , \quad t' = d-1 \quad , \quad s' = r \quad .$$

Consequently, a linear tight design (= generalized Hadamard code) is nothing but the dual of a linear perfect code (that is, a Hamming code, a Golay code, or a binary repetition code of odd length).

Finally, let us mention a criterion for a linear code Y having s weights to carry an association scheme with s classes. By definition, the row B'(x) of the outer distribution matrix B' of Y' is the weight distribution of the coset code x+Y', for any given $x \in X = F^n$. We shall denote by s^* the number of distinct weight distributions B'(x), with $x \notin Y'$. It follows from theorem 4 that the degree s of Y (i.e. the external distance r' of Y') is equal to rank(B')-1. Hence, $s \leq s^*$ holds. It turns out that we have $s = s^*$ if and only if the code Y carries a subscheme, with s classes, of the Hamming scheme H(n,q). Moreover, this subscheme admits a dual (cf. section 1), which has a natural representation on the cosets of Y' in X: two cosets are associated according to the weight distribution of their difference.

REFERENCES

[1] BIGGS, N.L., *Perfect codes in graphs*, J. Combinatorial Theory B, 15 (1973) 289-296.

[2] BOSE, R.C., *Strongly regular graphs, partial geometries and partially balanced designs*, Pacific J. Math., 13 (1963) 389-419.

[3] BOSE, R.C. & D.M. MESNER, *On linear associative algebras corresponding to association schemes of partially balanced designs*, Ann. Math. Statist., 30 (1959) 21-38.

[4] BOSE R.C. & T. SHIMAMOTO, *Classification and analysis of partially balanced incomplete block designs with two associate classes*, J. Amer. Statist. Assoc., 47 (1952) 151-184.

[5] CAMERON, P.J., *Near-regularity conditions for designs*, Geometriae Dedicata (to appear).

[6] DELSARTE, P., *Weights of linear codes and strongly regular normed spaces*, Discrete Math., 3 (1972) 47-64.

[7] DELSARTE, P., *Four fundamental parameters of a code and their combinatorial significance*, Information and Control, 23 (1973) 407-438.

[8] DELSARTE, P., *An algebraic approach to the association schemes of coding theory*, Philips Res. Repts. Suppl., 10 (1973).

[9] DOOB, M., *On graph products and association schemes*, Utilitas Math., 1 (1972) 291-302.

[10] GOETHALS, J.M. & J.J. SEIDEL, *Strongly regular graphs derived from combinatorial designs*, Canad. J. Math., 22 (1970) 597-614.

[11] GOETHALS, J.M. & S.L. SNOVER, *Nearly perfect binary codes*, Discrete Math., 3 (1972) 65-88.

[12] HANANI, H., *The existence and construction of balanced imcomplete block designs*, Ann. Math. Statist., 32 (1961) 361-386.

[13] HIGMAN, D.G., *Combinatorial considerations about permutation groups*, Lecture Notes, Mathematical Institute, Oxford, 1972.

[14] HUGHES, D.R., *Combinatorial analysis. t-designs and permutation groups*, Amer. Math. Soc., Proc. Symp. Pure Math., 6 (1962) 39-41.

[15] JOHNSON, S.M., *A new upper bound for error-correcting codes*, IEEE Trans. Information Theory, IT-8 (1962) 203-207.

[16] KRAWTCHOUK, M., *Sur une généralisation des polynômes d'Hermite*, Comptes Rendus de L'Académie des Sciences, Paris, 189 (1929) 620-622.

[17] LENSTRA, Jr., H.W., *Two theorems on perfect codes*, Discrete Math., 3 (1972) 125-132.

[18] LINT, J.H. VAN, *A survey of perfect codes*, Rocky Mountain J. Math. (to appear).

[19] MACWILLIAMS, F.J., Doctoral Dissertation, Harvard University, 1961 (unpublished).

[20] MACWILLIAMS, F.J., *A theorem on the distribution of weights in a systematic code*, Bell Syst. Tech. J., 42 (1963) 79-94.

[21] OGASAWARA, M., *A necessary condition for the existence of regular and symmetrical PBIB designs of* T_m *type*, Inst. Statist. mimeo series 418, Chapel Hill, N.C., 1965.

[22] OGAWA, J., *The theory of the association algebra and the relationship algebra of a partially balanced incomplete block design*, Inst. Statist. mimeo series 224, Chapel Hill, N.C., 1959.

[23] RAO, C.R., *Factorial experiments derivable from combinatorial arrangements of arrays*, J. Roy. Statist. Soc., 9 (1947) 128-139.

[24] SEMAKOV, N.V., V.A. ZINOV'EV & G.V. ZAITZEV, *Uniformly packed codes*, Problemy Peredači Informacii, 7 (1971) 38-50, (in Russian).

[25] TAMASCHKE, O., *Zur Theorie der Permutationsgruppen mit regulärer Untergruppe, I and II*, Math. Z., 80 (1963) 328-352 and 443-465.

[26] WILSON, R.M., *Lectures on t-designs*, Ohio State University, 1971, communicated by J. DOYEN.

[27] WILSON, R.M. & D.K. RAY-CHAUDHURI, *Generalization of Fisher's inequality to t-designs*, Amer. Math. Soc. Notices, 18 (1971) 805.

[28] YAMAMOTO, S., Y. FUJII & N. HAMADA, *Composition of some series of association algebras*, J. Sci. Hiroshima Univ., (A-I) 29 (1965) 181-215.

MATHEMATICAL CENTRE TRACTS 55, 1974, 162-182

RECENT RESULTS ON PERFECT CODES AND RELATED TOPICS

J.H. VAN LINT

Technological University, Eindhoven, The Netherlands

1. INTRODUCTION

In this paper we shall use the framework and terminology explained by DELSARTE in the previous paper [3]. We consider perfect codes in the Hamming and Johnson schemes and in association schemes corresponding to distance-transitive graphs. For these schemes we illustrate some recent examples and concentrate on the recent developments concerning non-existence proofs. Here the main tools are the *sphere packing bound* (cf. [3, § 4.1]) and *Lloyd's theorem* [3, Theorem 7].

The completely regular codes briefly mentioned by DELSARTE contain two classes which have properties similar to the perfect codes. The most interesting of these properties is that such codes can be used to construct t-designs. The search for such codes started a few years ago. Again we shall illustrate some examples and prove some non-existence theorems.

2. HAMMING SCHEMES H(n,q) WITH q A PRIME POWER

Let $q = p^r$ (p a prime). We consider a perfect e-error-correcting code Y, i.e. a perfect code of order e. The minimum distance of Y is $d = 2e+1$. For $e = 1$ there are many known examples of perfect codes (cf. [7]). For $e > 1$ one always has the trivial example $e = n$ and $|Y| = 1$. For $q = 2$ and $n = 2e+1$ the repetition code $Y := \{(0,0,\ldots,0),(1,1,\ldots,1)\}$ provides an example. Besides these there are 2 non-trivial perfect codes known as the Golay codes (cf. [7]). The parameters of these codes are $n = 23$, $q = 2$,

$e = 3$ respectively $n = 11$, $q = 3$, $e = 2$. For both of these codes it was shown recently [4], [12] that they are unique (up to translations and permutations of coordinate places). In 1970 VAN LINT [8] proved that if there are any other perfect codes then they have $e > 3$ and $p \mid e$. A year later TIETÄVÄINEN [13] proved that if $e \geq 3$ and $p \leq e$ then there is no perfect code of order e in $H(n,q)$, thus completely settling the problem for the case where q is a prime power. The same result was obtained independently by ZINOV'EV & LEONT'EV [15]. All these theorems had quite complicated proofs. A few months ago TIETÄVÄINEN [14] succeeded in shortening the proof considerably for $q > 2$. It turns out that very little more is needed if $q = 2$. We shall present the complete proof below.

We remind the reader of the definition of the *Krawtchouk polynomial* $K_k(n,q;u)$ of degree k:

$$K_k(n,q;u) := \sum_{j=0}^{k} (-1)^j (q-1)^{k-j} \binom{u}{j} \binom{n-u}{k-j} . \tag{2.1}$$

The sum polynomial Ψ_e occurring in Lloyd's theorem which we quote below is

$$\Psi_e(x) := K_e(n-1,q;x-1) = \sum_{i=0}^{e} (-1)^i \binom{n-x}{e-i} \binom{x-1}{i} (q-1)^{e-i} . \tag{2.2}$$

The two necessary conditions for the existence of a perfect code of order e in $H(n,q)$ mentioned in [3] are

$$\sum_{i=0}^{e} \binom{n}{i} (q-1)^i = q^k \tag{2.3}$$

for some integer k (cf. [7]) and Theorem 7 of [3] which states

$$\begin{cases} \Psi_e \text{ has } e \text{ distinct zeros } x_1 < x_2 < \ldots < x_e \text{ which are integers} \\ \text{in } [1,n]. \end{cases} \tag{2.4}$$

The following properties of Ψ_e and its zeros are easily obtained by substitution or by calculating suitable coefficients of Ψ_e.

$$\Psi_e(0) = \sum_{i=0}^{e} \binom{n}{i} (q-1)^i , \tag{2.5}$$

$$\Psi_e(1) = \binom{n-1}{e} (q-1)^e , \tag{2.6}$$

$$(2.7)\qquad \sum_{i=1}^{e} x_i = \frac{e(n-e)(q-1)}{q} + \frac{e(e+1)}{2} ,$$

$$(2.8)\qquad \sum_{i=1}^{e} x_i \le \frac{ne(q-1)}{q} \qquad (\text{for } q > 2) ,$$

$$(2.9)\qquad \prod_{i=1}^{e} x_i = e!q^{-e}\psi_e(0) ,$$

$$(2.10)\qquad \prod_{i=1}^{e} (x_i-1) = e!q^{-e}\psi_e(1) .$$

For integral values of x between 1 and n the terms of the sum in (2.2) alternate in sign. Since the terms decrease in absolute value, with increasing i, for $x < \frac{(n-e+1)(q-1)+e}{q-1+e}$ we have (if ψ_e has integral zeros) :

$$(2.11)\qquad x_1 \ge \frac{(n-e+1)(q-1)+e}{q-1+e} .$$

It was suggested by D.H. SMITH that the following lemma could prove to be useful in non-existence proofs of perfect codes.

LEMMA 1. *If a non-trivial perfect code of order* e *in* H(n,q) *exists then*

(i) $n \ge \frac{1}{2}e^2 + \frac{5}{2}e + 1$ *if* $q > 2$,

(ii) $n \ge e^2 + 4e + 2$ *if* $q = 2$.

PROOF. Let Y be a non-trivial perfect code of order e and let $\underline{a} = (a_0, a_1, \ldots, a_n)$ be the distance distribution of Y. Then we have

$$a_{2e+1} = e!n!(q-1)^{e+1}\{(n-e-1)!(2e+1)!\}^{-1} ,$$

$$a_{2e+2} = \tfrac{1}{2}a_{2e+1}\{\frac{q-1}{e+1}(n-e^2-3e-1) + e\} ,$$

and

$$a_{2e+3} = a_{2e+1}(n-2e-1)(n-e^2-4e-2)\{(2e+2)(2e+3)\}^{-1} \text{ if } q = 2.$$

Then (i) and (ii) follow from the observation that $a_{2e+2} \ge 0$ and $a_{2e+3} \ge 0$. □

We need one more concept which will play an important role in the non-existence proof. We define for $n \in \mathbb{N}$

(2.12) $$a_p(n) := \max\{m \in \mathbb{N} \mid m|n,\ p\nmid m\} ,$$

i.e. $a_p(n)$ is the largest divisor of n which is not divisible by p.

We call n_1 and n_2 p-*equivalent* if $a_p(n_1) = a_p(n_2)$.

Since one can explicitly determine the zeros of Ψ_2 it is easy to show that the ternary Golay code is the only non-trivial perfect code of order 2 (cf. [7]). In the following theorem we therefore take e > 2.

THEOREM 1. *If* $q = p^r$, $e > 2$ *then there is no perfect code* Y *of order* e, *with* $|Y| > 2$, *in* H(n,q).

PROOF. Assume Y is a perfect code of order $e < (n-1)/2$ in H(n,q). For the zeros of Ψ_e we find from (2.3), (2.5) and (2.9)

(2.13) $$\prod_{i=1}^{e} x_i = e!q^{k-e} ,$$

and hence

$$a_p(x_1)a_p(x_2) \ldots a_p(x_e) = a_p(e!) \le e! .$$

It follows that there are zeros x_i, x_j which are p-equivalent or $\{a_p(x_1), a_p(x_2), \ldots, a_p(x_e)\} = \{1,2,\ldots,e\}$, i.e. $p > e \ge 3$. In the latter case there is a zero $x_i = p^\alpha$ and a zero $x_j = 2p^\beta$ and then either $x_i \ge 2x_j$ or $x_j \ge 2x_i$. Hence we always have

(2.14) $$2x_1 \le x_e \text{ and hence } x_1x_e \le \frac{2}{9}(x_1+x_e)^2 .$$

Now by (2.5), (2,8), (2.14) and the arithmetic-geometric mean inequality we find

(2.15) $$(q-1)^e q^{-e} n(n-1)\ldots(n-e+1) < e!q^{-e}\Psi_e(0) = \prod_{i=1}^{e} x_i \le$$

$$\le \frac{8}{9}\left(\frac{x_1+x_e}{2}\right)^2 \left(\frac{x_2+x_3+\ldots+x_{e-1}}{e-2}\right)^{e-2} \le$$

$$\le \frac{8}{9}\left(\frac{x_1+\ldots+x_e}{e}\right)^e \le \frac{8}{9}(q-1)^e q^{-e} n^e$$

(for q > 2).

If q = 2 the final expression is $\frac{8}{9}\left(\frac{n+1}{2}\right)^e$.

Now let $q > 2$. From (2.6) and (2.10) we find $q^e \mid (n-1)\ldots(n-e)$ and therefore

$$(2.16) \qquad n > p^{re-[e/p]-[e/p^2]-\ldots} > p^{e(r-\frac{1}{p-1})} \geq q^{\frac{1}{2}e} .$$

By (2.15) we have

$$1 - \frac{e(e-1)}{2n} < \frac{n(n-1)\ldots(n-e+1)}{n^e} \leq \frac{8}{9} ,$$

i.e.

$$n < \frac{9}{2}e(e-1)$$

and hence (2.16) implies

$$3^{\frac{1}{2}e} \leq q^{\frac{1}{2}e} < \frac{9}{2}e(e-1) \text{ , i.e. } e \leq 11.$$

It then follows that $q \leq 8$ and $n \leq 495$. In 1967 (cf. [9]) a computer search had found all solutions of (2.5) for $n \leq 1000$, $q \leq 100$, $e \leq 1000$. This yielded no new codes. Hence for $q > 2$ the proof is finished.

For the case $q = 2$ ($n \neq 2e+1$) we do not have an inequality of type (2.16). We now have to use lemma 1 to get a lower bound on n and generalize the method used above to obtain an upper bound. Starting from (2.14) one shows by induction that

$$\prod_{i=1}^{s} \xi_i \leq \left(\frac{8}{9}\right)^{s-1}\left[\frac{1}{s}\sum_{i=1}^{s} \xi_i\right]^s$$

if $\xi_1,\xi_2,\ldots,\xi_s$ are 2-equivalent. Then in the same way as above the arithmetic-geometric mean inequality yields

$$(2.17) \qquad \prod_{i=1}^{e} x_i \leq \left(\frac{8}{9}\right)^{e-m}\left[\frac{x_1+\ldots+x_e}{e}\right]^e$$

if $x_1,x_2,\ldots,x_e$ are divided over m equivalence classes under 2-equivalence. This means that if we can prove that $m \leq e-6$ then the analogue of (2.15) extended by (2.17) gives us

$$n(n-1)\ldots(n-e+1) < \left(\frac{8}{9}\right)^6 (n+1)^e ,$$

i.e. $n < e^2+e$, which contradicts inequality (ii) of lemma 1.

In the proof of the remaining step we let $p(x)$ denote the product of the odd integers $\leq x$. Then

$$a_2(x_1x_2\ldots x_e) = a_2(e!) \le p(e)[\tfrac{e}{2}]!2^{-[\frac{e}{4}]} <$$

$$< p(e)e^{[\frac{e}{2}]-5} \qquad \text{(for } e \ge 16\text{)} .$$

Furthermore,

$$p(e) \le p(2m)e^{[\frac{e+1}{2}]-m}$$

and

$$a_2(x_1x_2\ldots x_e) \ge 1.3.5\cdots(2m-1) = p(2m) .$$

Combining these inequalities we find

$$1 < e^{[\frac{e}{2}]+[\frac{e+1}{2}]-m-5} ,$$

i.e. $m \le e-6$ (for $e \ge 16$). Hence the proof is complete for $e \ge 16$.

This leaves $e \le 15$ and then by (2.15) $n < 1000$ and we again refer to the computer search. □

Admittedly the case $q = 2$ is still rather messy. However, it seems likely that further simplifications of the proof are possible. We advise the reader to study the proof given here carefully in order to appreciate the great difficulties that arise when one tries to generalize to values of q which are not prime powers. In the next section we shall see that even the case $e = 2$, where one can explicitly determine the zeros of Ψ_e, is difficult.

3. HAMMING SCHEMES H(n,q) WITH q NOT A PRIME POWER

If q is not a prime power the sphere packing condition no longer has the form (2.3) which is replaced by

$$(3.1) \qquad \sum_{i=0}^{e} \binom{n}{i}(q-1)^i \mid q^n .$$

The other necessary condition for the existence of a perfect code still has the form (2.4).

Condition (3.1) is satisfied for e = 1 and n = q+1. It has been shown that a perfect code of order 1 in H(7,6) does not exist. As far as we know this is the only case where non-existence has been proved (for q not a prime power). Attempts to generalize the non-existence proofs for $e \geq 2$ have failed up to now because (3.1) is so much weaker than (2.3) that as a consequence (2.13) is replaced by a weaker statement. But even if (2.13) remained true the idea of splitting the zeros of Ψ_e into equivalence classes, which was the essential step in section 2, cannot be generalized.

As a small step on the road to complete understanding of perfect codes we shall completely treat the case q = 10, e = 2 since the alphabet of 10 symbols is of practical interest and this case illustrates how some of the ideas of section 2 can still be used.

We assume that a perfect code of order e in H(n,q) exists. From now on we take q = 10 but continue to use the symbol q in view of application to other examples. We shall keep e arbitrary as long as possible and then specialize to e = 2. The sphere packing bound now reads

$$\sum_{i=0}^{e} \binom{n}{i}(q-1)^i = q^k p^\alpha \ , \tag{3.2}$$

where p = 2 or p = 5. We define $\bar{p}$ by $q = p\bar{p}$. Since $\sum_{i=0}^{n} \binom{n}{i}(q-1)^i = q^n$ we find by subtraction that

$$q^k p^\alpha (q^{n-k-\alpha}\, \bar{p}^{\alpha} - 1) \equiv 0 \pmod{(q-1)^e} \ .$$

For $e \geq 2$ this implies that $\alpha \equiv 0 \pmod 6$. For the zeros of Ψ_e we again have (2.8) and (2.11). Instead of (2.9) we have

$$\prod_{i=1}^{e} x_i = e!\, q^{k-e} p^\alpha \ . \tag{3.3}$$

Furthermore we find from (2.6) and (2.10) in precisely the same way as (2.16) the inequality

$$n > 5^{\frac{3}{4}e} \tag{3.4}$$

(from the fact that 5^e divides (n-1)...(n-e)).

Next we remark that the argument of (2.15) still holds if $2x_1 \leq x_e$. This would again yield the inequality $n < \frac{9}{2}e(e-1)$, which contradicts (3.4). Hence we now have

(3.5) $$2x_1 > x_e \ .$$

THEOREM 2. *There is no perfect code of order* 2 *in* H(n,10) *for* n > 2.

PROOF. Assume on the contrary that such a code exists. By the sphere packing bound we have

(3.6) $$(18n-7)^2 - 41 = 8q^k p^\alpha \ .$$

Since $29^2 - 41 = 8q^2$ we find by subtraction

(3.7) $$36(9n+11)(n-2) = 8(q^k p^\alpha - q^2) \ .$$

Lloyd's theorem states that the zeros x_1, x_2 of

(3.8) $$x^2 - \{\frac{9}{5}(n-2)+3\}x + 2q^{k-2}p^\alpha$$

are integers between 1 and n.

We can already draw the following conclusions:

(3.9) $$n \equiv 2 \pmod 5 \ ,$$

(3.10) $$k \geq 2 \ ,$$

the latter because $x_1 x_2$ is an integer only if $k \geq 2$ or $k = 1$ and $p = 5$ but in that case $x_1 x_2 = 5^{\alpha-1}$ which contradicts (3.5). From (3.9), (3.10) and (3.7) we find that $n \equiv 2 \pmod{25}$ and using this we see from (3.8) that $x_1 + x_2 \not\equiv 0 \pmod 5$. It follows that one of the zeros is a power of 2. Let

$$x_1 = 2^\nu \ , \ x_2 = 2^\mu 5^\sigma$$

(where we no longer require x_1 to be the smaller of the two zeros).
We consider the equation (3.7) mod 32. The right-hand side is 0 and therefore we have two possibilities to consider, namely $n \equiv 2 \pmod 8$ and $n \equiv 5 \pmod 8$.
Case i : $n \equiv 2 \pmod 8$. In (3.8) one of the zeros is odd and since $x_1 \neq 1$ we must have $\mu = 0$. Furthermore, $x_1 + x_2$ is divisible by 3 which implies that $\nu + \sigma$ is odd. We now return to the sphere packing bound (3.6) with the knowledge that

$$\frac{9}{5}(n-2) + 3 = 2^\nu + 5^\sigma \ (\nu+\sigma \text{ odd}) \ .$$

Substitution yields

$$(5\cdot 2^{\nu+1}+2\cdot 5^{\sigma+1}-1)^2 - 41 = 8\cdot 10^k\cdot p^\alpha ,$$

i.e.

$$2^2\cdot 5^{2\sigma+2} - 2^2\cdot 5^{\sigma+1} \equiv 40 \pmod{32},$$

which is a contradiction.

Case ii : $n \equiv 5 \pmod 8$. We now have $x_1 + x_2 \equiv 2 \pmod 8$ and since (2.11) implies $x_1 \neq 2$ we must have $x_2 = 2\cdot 5^\sigma$. Again we substitute in (3.6). We find

$$-5\cdot 2^{\nu+2} \equiv 40 \pmod{25}, \text{ i.e. } \nu \equiv 3 \pmod 4$$

and

$$2^4\cdot 5^{2\sigma+2} - 2^3\cdot 5^{\sigma+1} \equiv 40 \pmod{64} \text{ unless } k = 2 \text{ and } p = 5 .$$

If $k = 2$ and $p = 5$ then $x_1 = 16$ and hence by (2.11) $n \leq 20$ contradicting $n \equiv 2 \pmod{25}$. So we must have σ even. Then

$$x_1 + x_2 = 2^\nu + 2\cdot 5^\sigma \equiv 1 \pmod 3$$

which contradicts (3.8).

This completes the proof. □

Of course this is an isolated example of a non-existence proof. In order to generalize this, e.g. to all q which are twice a prime, more ideas are necessary. We hope that some of the ideas used above will prove fruitful in future research on perfect codes.

4. JOHNSON SCHEMES J(n,v)

Let $n = 2e+1$ and $v = 2n$. We consider the two word code $Y := \{(1,1,\ldots,1,0,0,\ldots,0),(0,0,\ldots,0,1,1,\ldots,1)\}$ which we can interpret as an analogue of the repetition code in a Hamming scheme. Clearly every element of J(n,v) has distance $\leq e$ to exactly one of the elements of Y, i.e. Y is a perfect code of order e. These examples and the perfect codes with $|Y| = 1$ and $e = v$ we again consider trivial. No example of a non-trivial perfect code in J(n,v) is known. In a search for such codes one

quickly sees that it is again the sphere packing bound which is difficult to exploit. We shall briefly illustrate the case e = 2. We then have the two necessary conditions for the existence of a perfect code of order 2 in J(n,v):

$$(4.1) \qquad \{1 + \binom{n}{1}\binom{v-n}{1} + \binom{n}{2}\binom{v-n}{2}\} \mid \binom{v}{n} ,$$

$$(4.2) \qquad 4\Psi_2(x) = x^2 + \{2n^2 - 2vn + v - 6\}x +$$

$$+ \{n^4 - 2vn^3 + (v^2+v-5)n^2 + (-v^2+5v)n + 4\}$$

has two zeros which are both integers of the form i(v+1-i) with $0 \le i \le n$. So far the only solutions to these two conditions which we have been able to find correspond to trivial codes.

5. OTHER METRIC SCHEMES; GRAPHS

Let X be the set of rowvectors with 7 coordinates, three of which are 1 and the others 0. For $\underline{x},\underline{y} \in X$ we define

$$\rho(\underline{x},\underline{y}) = 1 \text{ if } (\underline{x},\underline{y}) = 0 ,$$

$$\rho(\underline{x},\underline{y}) = 2 \text{ if } (\underline{x},\underline{y}) = 2 ,$$

$$\rho(\underline{x},\underline{y}) = 3 \text{ if } (\underline{x},\underline{y}) = 1 ,$$

where $(\underline{x},\underline{y})$ is the inner product over $\mathbb{R}$. It is easy to check that this is a distance function for X. With each vector in X we associate a vertex $v(\underline{x})$ of a graph G and we join the vertices $v(\underline{x})$ and $v(\underline{y})$ by an edge iff $\rho(\underline{x},\underline{y}) = 1$. It turns out that $\rho(\underline{x},\underline{y})$ is the distance of $v(\underline{x})$ and $v(\underline{y})$ in the graph G. It is straightforward to check that the distance ρ defines a metric scheme in the sense of Remark (ii) of [3, §2].

G is a perfectly regular graph with valency $v_1 = 4$. Let Y be the set of 7 rowvectors of the incidence matrix of PG(2,2). If $\underline{x}$ and $\underline{y}$ are two distinct rows of Y then clearly $\rho(\underline{x},\underline{y}) = 3$. Since $|X| = 35$ we have

$$(1+v_1)\ |Y| = |X| ,$$

i.e. equality in the sphere packing bound [3, § 4.1, formula (3)]. Hence Y is a perfect code of order 1 in the metric scheme. This is the first example given by BIGGS [1] in his paper on perfect codes in distance-

transitive graphs. In [2] a number of other examples is given, all with e = 1. We have illustrated the example here in the setting of [3]. In both points of view it is clear that proving that Y is a perfect code is easy compared to showing that we have a metric scheme to start with. In the terminology of graph theory the difficult problem is to show that G is a distance-transitive graph and not to find the perfect code. Theorems of the type we discussed for the Hamming schemes were possible because we had an infinite class of schemes in which we could search for perfect codes. It does not seem likely that this will be the case for distance-transitive graphs. So even though we still have Lloyd's theorem as a tool we do not know where to use it. It would be extremely interesting if a perfect code of order e > 1 would be found in a scheme of the type considered here. Of course the Golay code is such a code and there is a code derived from the Golay code which is also of order 3 (O. HEDEN, private communication) but this code is essentially the same as the Golay code. An example not corresponding to a Hamming scheme is not known.

6. NEARLY PERFECT CODES

In this section we discuss a class of completely regular codes namely the nearly perfect codes introduced by GOETHALS & SNOVER [5].

JOHNSON [6] proved the following extension of the sphere packing bound (see also section 7):

LEMMA 2. *If* Y *is a code with minimum distance* d = 2e+1 *in* X := H(n,2) *then*

$$(6.1) \qquad |Y| \left\{ \sum_{i=0}^{e} \binom{n}{i} + \frac{1}{[n/(e+1)]} \binom{n}{e} \left(\frac{n-e}{e+1} - \left[\frac{n-e}{e+1} \right] \right) \right\} \leq 2^n = |X| \ .$$

If e+1 divides n+1 then this reduces to the sphere packing bound. It is well known that (e+1)|(n+1) is a necessary condition for the existence of a perfect code in H(n,2). The code Y is called *nearly perfect* if equality holds in (6.1). From the proof of (6.1) it immediately follows that if Y is a nearly perfect code then for every $\underline{x} \in X$ with $\rho(\underline{x},Y) > e$ there are exactly [n/(e+1)] points $\underline{y} \in Y$ with $\rho(\underline{x},\underline{y}) = e+1$. Furthermore, it follows that if $\rho(\underline{x},\underline{y}) = e$ for some $\underline{y} \in Y$, then there are exactly [(n-e)/(e+1)] points $\underline{z} \in Y$ with $\rho(\underline{x},\underline{z}) = e+1$. In fact, such a code Y is completely regular. The distance distribution of Y is determined in [5].

The following theorem is an example of the theorems given in [5] showing the importance of nearly perfect codes in the theory of designs. We shall interpret a point of Y as the incidence vector of a subset of a set S of n points.

THEOREM 3. *If* Y *is a nearly perfect code with minimum distance* d = 2e+1 *in* X = H(n,2) *and* $\underline{0} \in$ Y *then the words of weight* d *in* Y *form an* e*-design with* $\lambda = [(n-e)/(e+1)]$.

PROOF. Any e-subset D of S corresponds to a point $\underline{x} \in$ X with weight e, i.e. distance e to $\underline{0}$. We mentioned above that $\underline{x}$ then has distance e+1 to exactly λ points of Y each of which therefore has weight 2e+1. Hence D is a subset of exactly λ sets corresponding to code words of weight 2e+1. □

It is also shown in [5] that such designs can be extended to (e+1)-designs.

The non-linear codes known as the Preparata codes (cf. [10]) have $n = 4^m-1$, $|Y| = 2^{n-r}$ where r = 4m-1 (m ≥ 2), and d = 5. By substitution we see that these codes satisfy (6.1) with equality, i.e. they are nearly perfect. We thus obtain an infinite class of 3-designs.

Of course the question now rises whether nearly perfect codes with e > 2 can be found. (From now on we exclude perfect codes.) The definition alone is enough to show that the answer is negative for e = 3 and e = 4. GOETHALS & SNOVER mention this in their paper without giving the proof. We present their proof here.

THEOREM 4. *Except for the known perfect codes there are no nearly perfect codes with minimum distance* 7 *or* 9.

PROOF.

(i) Suppose $n \not\equiv 3 \pmod 4$. Let $n+1 \equiv s \pmod 4$ where s = 1,2 or 3.

Then substituting this and e = 3 in (6.1) yields

$$|Y|\{1 + n + \binom{n}{2} + \frac{n+1}{n+1-s}\binom{n}{3}\} = 2^n \quad . \tag{6.2}$$

This can be written as

$$|Y|\{6 + (n+1)(n^2+(s-1)n+(s-2)(s-3))\} = 3\cdot 2^{n+1} \quad . \tag{6.3}$$

If s = 1, i.e. $n \equiv 0 \pmod 4$, we find from (6.3)

$$|Y|(n+2)(n^2-n+4) = 3\cdot 2^{n+1}$$

and hence $(n+2)|6$ which gives us $n = 4$, which does not correspond to a nearly perfect code. If $s = 2$ or 3 we obtain from (6.3)

$$|Y|(n+1)n(n+s-1) = 6\{2^n - |Y|\} ,$$

where $|Y| = 2^k$ or $3\cdot 2^k$ $(k < n)$. This implies that

$$(n+1)n(n+s-1) = 6(2^{n-k}-1) \text{ or } 2(2^{n-k}-3) .$$

Here the left-hand side is $\equiv 0 \pmod 4$ and the right-hand side is $\equiv 2 \pmod 4$, a contradiction.

(ii) We now consider the case $e = 4$. Let $n+1 \equiv s \pmod 5$, where $s = 1,2,3$ or 4. Again we substitute in (6.1) and replace $n+1$ by m. We find

$$3\cdot 2^{n+3}/|Y| = \begin{cases} m\,(m^3-5m^2+14m+8) & \text{if } s = 1 , \\ m\,(m^3-4m^2+7m+20) & \text{if } s = 2 , \\ m\,(m^3-3m^2+2m+24) & \text{if } s = 3 , \\ m\,(m^3-2m^2-m+26) & \text{if } s = 4 . \end{cases}$$

Clearly $16 \nmid m$ and therefore $m \mid 24$ which leaves only a finite number of cases which are all easily ruled out. □

Of course the first really interesting case is $e = 5$ since it could lead to a 6-design. In this case we were not able to do anything with (6.1) alone. However, GOETHALS & SNOVER also proved that there is an analogue of Lloyd's theorem for nearly perfect codes. We quote the theorem.

THEOREM 5. *Let there exist a nearly perfect code in* $H(n,2)$ *with minimum distance* $2e+1$ *and let* $n+1 \not\equiv 0 \pmod{e+1}$. *Then the polynomial*

$$(6.4) \qquad Q(x) := \Psi_{e-1}(x) + \frac{1}{[(n+1)/(e+1)]}\,(\Psi_{e+1}(x) - \Psi_{e-1}(x))$$

(where $\Psi_e(x)$ *is the polynomial of* (2.2) *with* $q = 2$*) has* $e+1$ *distinct integral zeros between* 1 *and* n.

As was to be expected the case $e = 5$ does not yield any solutions either.

THEOREM 6. *There is no nearly perfect code with minimum distance* 11 *in* H(n,2) *for* n > 11.

PROOF. Assume that Y is such a nearly perfect code. We write $n = 6v+\ell$ where $\ell = 0,1,2,3$ or 4 (since $\ell = 5$ is excluded by the theorem on perfect codes). By substitution in (6.1), taking e = 5, we see that $|Y|$ is either a power of 2 or 5 times a power of 2. Hence we have

$$(6.5) \qquad \sum_{i=0}^{5} \binom{n}{i} + \frac{1}{[n/6]} \binom{n}{5} \left\{ \frac{n-5}{6} - \left[\frac{n-5}{6}\right] \right\} = 2^r/a \ ,$$

where a = 1 or 5. As in the case of perfect codes the left-hand side of (6.5) is Q(0). By substitution we see that Q(1) > 0.

In the same way as (2.7) and (2.9) we find the sum and product of the zeros of Q:

$$(6.6) \qquad \sum_{i=1}^{6} x_i = 3(n+1) \ ,$$

$$(6.7) \qquad \prod_{i=1}^{6} x_i = \left[\frac{n+1}{6}\right] 6! \, 2^{r-6}/a \ .$$

We observe that Q(n+1-x) = Q(x), i.e.

$$x_{7-i} = n+1 - x_i \qquad (i=1,2,3).$$

We now introduce the variable $z := (2x-n-1)^2$. On substitution in (6.4) we then find that

$$(6.8) \qquad Q^*(z) := z^3 + 5(-2n+5-\ell)z^2 + \{15n^2 + 10(3\ell-5)n - (70\ell-19)\}z +$$
$$- 15(\ell+1)(n-1)(n-3)$$

has three integral zeros $z_i := (2x_i-n-1)^2$, i=1,2,3. Again a simple computation shows that $Q^*(0) < 0$ and $Q^*(8) > 0$. Hence by theorem 5 either n is even and $Q^*(1) = 0$ or n is odd and $Q^*(4) = 0$. However $Q^*(4) < 0$ for $\ell > 0$ and $Q^*(1) = -15\ell(n-2)(n-4) = 0$ only if $\ell = 0$. Hence we now know that n = 6v and that $x_3 = 3v$ and $x_4 = 3v+1$. Furthermore, z_1 and z_2 satisfy

$$z^2 + (26-10n)z + 15(n-1)(n-3) = 0 \ ,$$

i.e.

$$z_{1,2} = 5n - 13 \pm (10n^2-70n+124)^{\frac{1}{2}} = 5n - 13 \pm x \ ,$$

where

$$5(2n-7)^2 + 3 = 2x^2 \quad . \tag{6.9}$$

Substituting $x_3 = 3v$ and $x_4 = 3v+1$ in (6.7) we find

$$(3v+1)x_1x_2x_5x_6 = 3\cdot 5\cdot 2^{r-2}/a \ ,$$

so either $3v+1 = 2^\sigma$ or $3v+1 = 5\cdot 2^\sigma$, i.e. $n = 2(2^\sigma-1)$ or $n = 2(5\cdot 2^\sigma-1)$. First substitute $n = 2(2^\sigma-1)$ in (6.9). This yields

$$16(5\cdot 2^{2\sigma-1}-55\cdot 2^{\sigma-2}+19) = x^2 \ .$$

The expression in brackets is $\equiv 3 \pmod 4$ if $\sigma \geq 4$. Hence only $n = 14$ is a possibility, but this does not yield a solution. Substitution of $n = 2(5\cdot 2^\sigma-1)$ yields a contradiction in exactly the same way. □

7. UNIFORMLY PACKED CODES

The uniformly packed codes were introduced by SEMAKOV, ZINOV'EV & ZAITSEV [11]. Once again the codes are in $H(n,2)$. The definition generalizes the idea of perfect and nearly perfect codes (in fact these are uniformly packed).

Let Y be a binary code of length n and minimum distance d. Let $e := [(d-1)/2]$. The set of all words of length n is again denoted by X. We define

$$Y_e := \{\underline{x} \in X \mid \rho(\underline{x},Y) \geq e\} \ , \tag{7.1}$$

and

$$\forall \underline{z} \in Y_e, [r(\underline{z}) := |Y \cap S_{e+1}(\underline{z})|] \ , \tag{7.2}$$

i.e. $r(\underline{z})$ is the number of points $\underline{y} \in Y$ with $e \leq \rho(\underline{y},\underline{z}) \leq e+1$. Clearly we have

$$\forall \underline{z} \in Y_e, [r(\underline{z}) \leq [(n+1)/(e+1)] \ . \tag{7.3}$$

Since

$$\sum_{\underline{z}\in Y_e} r(\underline{z}) = \sum_{\underline{y}\in Y} |S_{e+1}(\underline{y}) \setminus S_{e-1}(\underline{y})| = |Y|(\binom{n}{e} + \binom{n}{e+1})$$

and

$$|Y_e| = 2^n - |Y| \sum_{i=0}^{e-1} \binom{n}{i}$$

we find that the average value r of $r(\underline{z})$ is

$$r = \frac{|Y| \binom{n+1}{e+1}}{2^n - |Y| \sum_{i=0}^{e-1} \binom{n}{i}} . \tag{7.4}$$

Observe that (7.3) and (7.4) together yield a proof of the Johnson bound (6.1).

If $\forall \underline{z} \in Y_e$, $[r(\underline{z}) = r]$ the code Y is called *uniformly packed*. Clearly a uniformly packed code with $r = [(n+1)/(e+1)]$ is nearly perfect (and of course perfect if $r = (n+1)/(e+1)$ and also if $r = 1$). From now on we only consider uniformly packed codes with $1 < r < [(n+1)/(e+1)]$. In [11] the distance distribution of a uniformly packed code is determined and it turns out that these codes are also completely regular. Furthermore, theorem 3 also generalizes. In fact, in the extended code of a uniformly packed code the words of a given weight w form an (e+1)-design. So once again the search starts!

We use the following notation

$$r = (n-s)/(e+1), \text{ where } s \geq e. \tag{7.5}$$

We restrict the search to $e = 1$, $s \leq 15$ and $e = 2$, $s \leq 5$.

From (7.4) we find

$$2^n(n-s) = |Y| \ (n^2+2n-s) . \tag{7.6}$$

Since

$$n^2 + 2n - s > (n-s)(n+s+2)$$

and

$$n + s + 2 \geq 2s + 6 \geq 8,$$

we see that $16 \mid (n^2+2n-s)$, i.e. $s \equiv 0,3,8$ or $15 \pmod{16}$.

(a) <u>e = 1, s = 3</u>. There are two cases to consider. If $3 \nmid n$ we have

$$(n+3)(n-1) = n^2 + 2n - 3 = 2^k ,$$

i.e. $n = 5$, contradicting $r > 1$. If $n = 3m$ we find

$$(3m-1)(m+1) = 2^k ,$$

i.e. $m = 1$ or $m = 3$. So we have a possible solution with $n = 9$ and $r = 3$. Then $|Y| = 2^5$. There is indeed a uniformly packed code with these parameters.

Let $C := \text{circulant}(0,1,1,0)$. Then the rows of

$$G := \left(I_5 \;\middle|\; \begin{matrix} 1 & 1 & 1 & 1 \\ & & & \\ & & C & \end{matrix} \right)$$

generate a 5-dimensional linear subspace Y of X. The rows of G have weight ≥ 3 (equality occurring) and no two rows of G coincide in the final 4 positions. Hence the non-zero elements of Y have weight ≥ 3. We know that the maximal value of $r(\underline{z})$ for $\underline{z} \in Y_1$ is at most 5 and the average value of $r(\underline{z})$ is 3. Suppose there is a $\underline{z} \in Y_1$ with $r(\underline{z}) = 5$. Then without loss of generality we may assume $w(\underline{z}) = 1$ and then the sum of the 4 points of Y at distance 2 from $\underline{z}$ would have to be $(1,1,\ldots,1) - \underline{z}$. Since $(1,1,\ldots,1) \in Y$ this would imply $\underline{z} \in Y$, a contradiction. If for some $\underline{z}$ we would have $r(\underline{z}) = 4$ then again without loss of generality we may take $w(\underline{z}) = 1$ or 2. If $w(\underline{z}) = 1$ then the sum of the 3 points of Y at distance 2 from $\underline{z}$ would have weight 7, again a contradiction. If $w(\underline{z}) = 2$ then the sum of the 4 points of Y at distance 2 from $\underline{z}$ would have weight 8, once more a contradiction. This proves that Y is uniformly packed. The 2-design formed by the words of weight 4 of the extended code $\bar{Y}$ is the residual of the symmetric (16,6,2)-design.

(b) $\underline{e = 1,\ s = 8}$. Now (7.6) reads

$$2^n(n-8) = |Y|(n^2+2n-8) = |Y|(n+4)(n-2) .$$

If $n \equiv 2 \pmod 3$ then $n = 9m+8$ and we find

$$m \cdot 2^n = |Y|(3m+4)(3m+2)$$

which implies $m = 0$. If $n \not\equiv 2 \pmod 3$ then $n-2 = 2^k$, $n+4 = 2^\ell$, i.e. $n = 4$, a contradiction.

(c) $\underline{e = 1,\ s = 15}$. In this case we find from (7.6)

$$(n-15)2^n = |Y|(n+5)(n-3) .$$

As before we distinguish 4 cases depending on the g.c.d. of n and 15. In the same way as above we then find two possible parameter sets for uniformly packed codes:

(7.7) $\qquad n = 27,\ r = 6,\ |Y| = 2^{21}$,

(7.8) $\qquad n = 35,\ r = 10,\ |Y| = 2^{29}$.

At present we do not know whether such codes exist.

We turn to the case e = 2. Then $n \equiv s \pmod 3$ and $n \geq s+6$. We find from (7.4) the equation

(7.9) $\qquad (n-s)2^{n+1} = |Y|(n+1)(n^2+n-2s)$

and we observe that

(7.10) $\qquad (n-s,n+1) = (s+1,n+1)$,

(7.11) $\qquad (n-s,\ n^2+n-2s) = (n-s,\ s(s-1))$.

(d) <u>e = 2, s = 2</u>. We then have from (7.9), (7.10) and (7.11)

$$n + 1 = 3\cdot 2^k ,$$

$$n^2 + n - 4 = 2^m .$$

If $k > 2$ then $m = 2$ which is impossible. If $k = 2$ we find $n = 11$ which yields a possible set of parameters: $n = 11$, $r = 3$, $|Y| = 24$. We now demonstrate a uniformly packed code with these parameters. Consider a Hadamard matrix H_{12} of order 12. From

$$A := \begin{pmatrix} \frac{1}{2}(H_{12} + J) \\ \frac{1}{2}(-H_{12} + J) \end{pmatrix}$$

we leave out the first column. The rows of the remaining matrix are the 24 words of the punctured Hadamard code Y. From the properties of the Hadamard matrix it follows that Y has minimum distance 5. For any $\underline{z} \in Y_2$ we know that $r(\underline{z}) \leq 4$ and the average value of $r(\underline{z})$ is 3. Suppose $r(\underline{z}) = 4$ for some $\underline{z}$. This implies that after a suitable multiplication of rows and columns by -1 and a permutation of columns there are 4 rows of H_{12} which have the form

x_1++ +++ +++ +++
x_2-- --- +++ +++
x_3-- +++ --- +++
x_4-- +++ +++ ---

(where in this notation $\underline{z}$ corresponds to (x-- +++ +++ +++)). Taking $x_1 = +1$ we must have $x_2 = x_3 = x_4 = -1$ and then there is no other row of 12 +1's and -1's which is orthogonal to these 4 rows, a contradiction. Hence there is no $\underline{z}$ with $r(\underline{z}) = 4$, and therefore the code Y is uniformly packed. If we extend Y we find A. The words of weight 6 are obtained by leaving out the first and thirteenth row of A (if A has standard form). This yields a well-known 3-design.

(e) $\underline{e = 2,\ s = 3}$. From (7.9), (7.10), (7.11) we find $n+1 = 2^k$. Since $n \equiv 0 \pmod 3$ k must be even and $k \geq 4$. Then

$$n^2 + n - 2s = 2^{2k} - 2^k - 6 = 3 \cdot 2^m ,$$

i.e. m = 1, a contradiction.

(f) $\underline{e = 2,\ s = 4}$ and $\underline{e = 2,\ s = 5}$ are treated in exactly the same way. We omit the details. No possible parameter sets come up.

The equation (7.4) has a number of infinite families of solutions. We mention one below, the others are still being investigated. Without going into details we mention that there is also a generalization of Lloyd's theorem for uniformly packed codes. In some cases the infinite families of solutions of (7.4) also satisfy the conditions of this theorem.

Let $k \geq 2$, $n = 2^{2k-1}-1$, $e = 2$, $r = \frac{1}{3}(4^{k-1}-1)$, $d = n-2(2k-1)$ and $|Y| = 2^d$. Then these numbers satisfy (7.4). For k = 2 these are the parameters of the repetition code (which is perfect). For $k \geq 2$ the parameters are those of the 2-error-correcting primitive binary BCH-codes of length n and dimension d (cf. [7]). These codes are indeed uniformly packed and therefore we find from these codes several infinite sequences of 3-designs. We leave the details for a later paper.

REFERENCES

[1] BIGGS, N.L., *Perfect codes in graphs*, J. Combinatorial Theory B, 15 (1973) 289-296.

[2] BIGGS, N.L., *Perfect codes and distance-transitive graphs*, in: *Combinatorics*, T.P. MCDONOUGH & V.C. MAVRON (eds.), Proc. of the Third British Comb. Conference, Aberystwyth 1973, London Math. Soc. Lecture Notes (1974).

[3] DELSARTE, P., *The association schemes of coding theory*, in: *Combinatorics*, M. HALL, JR. & J.H. VAN LINT (eds.), Mathematical Centre Tracts 55, Amsterdam, 1974, pp. 139-157.

[4] DELSARTE, P. & J.M. GOETHALS, *Unrestricted codes with the Golay parameters are unique*, Report R 238, M.B.L.E. Research Laboratory, Brussels, 1973.

[5] GOETHALS, J.M. & S.L. SNOVER, *Nearly perfect binary codes*, Discrete Math., 3 (1972) 65-88.

[6] JOHNSON, S.M., *A new upper bound for error-correcting codes*, IEEE Trans. Information Theory, IT-8 (1962) 203-207.

[7] LINT, J.H. VAN, *Coding theory*, Lecture Notes in Mathematics 201, Springer-Verlag, Berlin, 1971.

[8] LINT, J.H. VAN, *Nonexistence theorems for perfect error correcting codes*, in: *Computers in algebra and number theory*, SIAM-AMS Proceedings IV, Amer. Math. Soc., Providence, R.I., 1971, pp. 89-95.

[9] LINT, J.H. VAN, *1967-1969 Report of the discrete mathematics group*, Report 69-WSK-04 of the Technological University, Eindhoven, 1969.

[10] PREPARATA, F.P., *A class of optimum nonlinear double-error-correcting codes*, Information and Control, 13 (1968) 378-400.

[11] SEMAKOV, N.V., V.A. ZINOV'EV & G.V. ZAITSEV, *Uniformly packed codes*, Problems of Information Transmission, 7 (1971) 30-39 (translated from Problemy Peredači Informacii, 7 (1971) 38-50).

[12] SNOVER, S.L., *The uniqueness of the Nordstrom-Robinson and the Golay binary codes*, thesis, Michigan State Univ., 1973.

[13] TIETÄVÄINEN, A., *On the nonexistence of perfect codes over finite fields*, SIAM J. Appl. Math., 24 (1973) 88-96.

[14] TIETÄVÄINEN, A., *A short proof for the nonexistence of unknown perfect codes*, Ann. Acad. S . Fenn. A580, 1974.

[15] ZINOV'EV, V.A. & V.K. LEONT'EV, *A theorem on the nonexistence of perfect codes over finite fields*, Problemy Peredači Informacii, to appear (in Russian).

MATHEMATICAL CENTRE TRACTS 55, 1974, 183-200

IRREDUCIBLE CYCLIC CODES AND GAUSS SUMS *)

R.J. McELIECE

Jet Propulsion Labaratory, California Institute of Technology, Pasadena, Cal. 91109, USA

1. INTRODUCTION

In this paper we wish to point out the existence of a close connection between *irreducible cyclic codes* and *Gauss sums* over finite fields, and then to apply the well-developed theory of Gauss sums to the much less well-developed theory of irreducible cyclic codes.

We begin by giving two equivalent definitions of an irreducible cyclic code.

Let p be a prime, and let $q = p^e$ be a power of p. Denote by F_e the finite Galois field GF(q). Let n be a positive integer not divisible by p, and let $h(x) = h_0+h_1x+\ldots+h_kx^k$ be an F_e-irreducible divisor of $f_n(x)$, the n-th cyclotomic polynomial. It follows from the theory of finite fields that k, the degree of h(x), is the order of q mod n, i.e., the least positive integer such that $q^k \equiv 1 \pmod n$. The set of n-tuples $(c_0,c_1\ldots,c_{n-1})$ from F_e such that

$$\sum_{i=0}^{k} h_i c_{i+t} = 0 , \qquad (t=0,1,\ldots,n-1), \tag{1.1}$$

(subscripts are to be reduced mod n if necessary) is called an *(n,k) irreducible cyclic code* over F_e; h(x) is called the *parity-check polynomial* of the code.

Alternatively, if θ is a zero of h(x) in the field $F_{ek} = GF(q^k)$, the code can be characterized as the set of n-tuples $c(x) = (c_0(x),c_1(x),\ldots,c_{n-1}(x))$ from F_e of the form

---

*) This paper presents the results of one phase of research carried out at the Jet Propulsion Laboratory, California Institute of Technology, under Contract No. NAS 7-100, sponsored by the National Aeronautics and Space Administration.

(1.2) $$c_i(x) = T_e^{ek}(x\,\theta^i) \,, \qquad (i=0,1,\ldots,n-1) \,,$$

for some $x \in F_{ek}$, where $T_e^{ek}(\cdot)$ is the trace of F_{ek} over F_e. VAN LINT [10,Ch.3] gives a proof of the equivalence of these two formulations but for our purposes we shall take (1.2) as the definition of an irreducible cyclic code.

Now for $x \in F_{ek}$, $a \in F_e$, we denote by $\nu(x;a)$ the number of integers $i \in \{0,1,\ldots,n-1\}$ such that $c_i(x) = a$, i.e.,

(1.3) $$\nu(x;a) = \left|\{i: T_e^{ek}(x\,\theta^i) = a,\ 0 \le i \le n-1\}\right| \,.$$ *)

It is the study of the numbers $\nu(x;a)$ that interests us, but after section 2 we will consider only the numbers $\nu(x;0)$. Since $n-\nu(x;0)$ is the weight of $c(x)$, i.e., the number of non-zero components of $c(x)$, this restriction is equivalent to the study of the *weight distributions* of irreducible cyclic codes. It is probable, however, that many of the techniques developed in this paper can be extended to the numbers $\nu(x;a)$ for $a \neq 0$.

We will show in section 2 that the numbers $\nu(x;a)$ are intimately related to certain Gauss sums in the field F_{ek}. If μ is a character (a homomorphism into the complex numbers) of the multiplicative group $F_{ek}^{(\cdot)}$ of F_{ek} and λ is a character of the additive group $F_{ek}^{(+)}$ of F_{ek}, the Gauss sum $G(\mu,\lambda)$ is defined by

(1.4) $$G(\mu,\lambda) = \sum_{x \in F_{ek}^*} \mu(x)\lambda(x) \,.$$

We shall see in section 2 that the numbers $\nu(x;a)$ can be expressed in terms of the Gauss sums $G(\mu,\lambda)$ for characters μ of order $N = (q^k-1)/n$, i.e., characters satisfying $\mu^N(x) = 1$ for all $x \neq 0$. The values $\nu(x;0)$ actually only depend on the sums $G(\mu,\lambda)$ for which $\mu^{N_1} = 1$, where $N_1 = g.c.d.(N,(q^k-1)/(q-1))$. By invoking known theorems on Gauss sums we shall succeed in computing the numbers $\nu(x;0)$ for all irreducible cyclic codes for

*) Although the definition (1.2) depends upon the particular n-th root of unity θ, the replacement of θ by another such primitive n-th root of unity θ^j with $(j,n) = 1$ will only permute the coordinates of $c(x)$ and so will not affect the numbers $\nu(x;a)$.

which: $N_1 = 1$ (section 2), $N_1 = 2$ (section 5), $N_1 = 3$ (section 6), $N_1 = 4$ (section 7), $p^l \equiv -1 \pmod{N_1}$ for some l (section 3), $\mathrm{ord}_p(N_1) = (N_1-1)/2$ and N_1 is prime (section 4). We have collected the necessary facts about Gauss sums in the appendix.

Some of the results in this paper are already known, at least for $e = 1$. In particular, MCELIECE & RUMSEY [11] first noticed that the Davenport-Hasse theorem (G5) about Gauss sums could be applied to the study of irreducible cyclic codes. BAUMERT & MCELIECE [1] calculated $\nu(x;a)$ for all a if $p^l \equiv -1 \pmod{N}$ for some l, or if $N = 2$, and for $q = 2$ and all $N < 100$. BAUMERT & MYKKELTVEIT [2] calculated $\nu(x;a)$ for all a when N is a prime for which p generates the quadratic residues. Later, in unpublished manuscripts, MYKKELTVEIT settled the cases $N = 3$ and $N = 4$. The main contributions of this paper are the extension of previous results to $GF(q)$, and the observation that if one is only interested in the values $\nu(x;0)$, it is the number N_1 rather than N which is important.

2. GENERAL RESULTS

Let β be a complex primitive N-th root of unity, and ψ a primitive root in F_{ek} such that $\psi^N = \theta$. Then the function μ defined by

$$\mu(\psi^i) = \beta^i \tag{2.1}$$

is a character of order N of the multiplicative group $F_{ek}^{(\cdot)}$. Similarly if ζ is a complex primitive p-th root of unity the function λ defined by

$$\lambda(x) = \zeta^{T_1^{ek}(x)} \tag{2.2}$$

is a character of order p of the additive group $F_{ek}^{(+)}$. Now it is easily shown that an element $y \in F_{ek}^*$ is of the form $x\theta^j$ for some j if and only if $\mu^i(x) = \mu^i(y)$ for $i=0,1,\ldots,N-1$. Hence $\nu(x;a)$ is equal to the number of elements $y \in F_{ek}^*$ such that

$$\mu^i(yx^{-1}) = 1 , \qquad (i=0,1,\ldots,N), \tag{2.3}$$

$$T_e^{ek}(y) = a. \tag{2.4}$$

If ξ is a fixed element of F_{ek} with $T_e^{ek}(\xi) = 1$, condition (2.4) becomes $T_e^{ek}(y-a\xi) = 0$. If we define the character λ_b for $b \in F_{ek}$ by

$$\lambda_b(x) = \lambda(bx) \ , \tag{2.5}$$

it follows that (2.4) is equivalent to

$$\lambda_b(y-\xi a) = 1 \quad \text{for all } b \in F_e. \tag{2.4'}$$

Thus it follows from the fact

$$\sum_{i=0}^{N-1} \mu^i(x) = \begin{cases} N & \text{if } \mu(x) = 1, \\ 0 & \text{if } \mu(x) \neq 1, \end{cases} \qquad \sum_{b \in F_e} \lambda_b(x) = \begin{cases} q & \text{if } \lambda_b(x) = 1 \ \forall_b \in F_e, \\ 0 & \text{otherwise,} \end{cases}$$

that

$$qN\nu(x;a) = \sum_{y \in F_{ek}^*} \sum_{i=0}^{N-1} \sum_{b \in F_e} \mu^i(x^{-1}y)\lambda_b(y-a\xi) =$$

$$= \sum_{i=0}^{N-1} \mu^{-i}(x) \sum_{b \in F_e} \lambda_b^{-1}(a\xi) \sum_{y \in F_{ek}^*} \mu^i(y)\lambda_b(y) \ . \tag{2.6}$$

Now the sum $\sum \mu^i(y)\lambda_b(y)$ which appears in (2.6) is just the *Gauss sum* $G(\mu^i,\lambda_b)$ over the field F_{ek}. We begin our simplification of (2.6) by separating out the sums $G(\mu^i,\lambda_b)$ for which either $i = 0$ or $b = 0$, using property (G1). The result is

$$qN\nu(x;a) = \begin{cases} q^k - q + \sum_{i=1}^{N-1} \sum_{b \in F_e^*} \mu^{-i}(x)\lambda_b^{-1}(a\xi)\ G(\mu^i,\lambda_b) & \text{if } a = 0 \ , \\ q^k + \sum_{i=1}^{N-1} \sum_{b \in F_e^*} \mu^{-i}(x)\lambda_b^{-1}(a\xi)\ G(\mu^i,\lambda_b) & \text{if } a \neq 0 \ . \end{cases}$$

By (G2), if $b \neq 0$, $G(\mu^i,\lambda_b) = \mu(b^{-1})G(\mu^i,\lambda)$. Thus if we denote $G(\mu^i,\lambda)$ by G_i and $\mu(x)$ by β^j,

$$(2.7)\qquad q\mathrm{N}\nu(x;a) = \begin{cases} q^k - q + \sum_{i=1}^{N-1} \beta^{-ij} G_i \sum_{b\in F_e^*} \mu(b)^{-i} \lambda_{a\xi}(b)^{-1} & \text{if } a = 0, \\ q^k + \sum_{i=1}^{N-1} \beta^{-ij} G_i \sum_{b\in F_e^*} \mu(b)^{-i} \lambda_{a\xi}(b)^{-1} & \text{if } a \neq 0. \end{cases}$$

The inner sum in (2.7) is the complex conjugate $\overline{g(\mu^i, \lambda_{a\xi})}$ of the Gauss sum of μ^i and $\lambda_{a\xi}$ over the subfield F_e of F_{ek}. Now the character μ^i will be identically equal to 1 on F_e^* if and only if it is equal to 1 on a primitive root of F_e^*. But $N_e^{ek}(\psi) = \psi^{(q^k-1)/(q-1)} = g$ is such a primitive root, and so $\mu^i(g) = \beta^{i(q^k-1)/(q-1)}$. It follows that the character μ^i acts trivially on F_e^* if and only if $i \equiv 0 \pmod{N_2}$, where

$$(2.8)\qquad \begin{cases} N_1 = (N, (q^k-1)/(q-1)) \\ N_2 = N/N_1 \ . \end{cases}$$

We now distinguish two cases, $a = 0$ and $a \neq 0$. If $a = 0$, the inner sum in (2.7) is simply $\sum_{b\in F_e^*} \mu(b)^{-i}$, which is by our above remarks $q-1$ if $i \equiv 0 \pmod{N_2}$ and 0 otherwise. Hence (2.7) becomes in this case

$$(2.9)\qquad \nu(x;0) = \frac{q^{k-1}-1}{N} + \frac{(q-1)}{qN} \sum_{i=1}^{N_1-1} G_{N_2 i} \beta^{-N_2 ij} \ .$$

Using property (G3), we obtain the estimate

$$(2.10)\qquad \left| \nu(x;0) - \frac{q^{k-1}-1}{N} \right| \leq \frac{(q-1)(N_1-1)}{N} q^{k/2-1} \ .$$

In particular if $N_1 = 1$ (i.e., if $q \equiv 1 \pmod{N}$ and $(N,k) = 1$) it follows that

$$\nu(x;0) = \frac{q^{k-1}-1}{N} \qquad \text{for all } x \in F_{ek}^* \ ,$$

a result recently proved for $e = 1$ by OGANESIAN & YADZJAN [12].

Next consider the case $a \neq 0$. According to (G2), $g(\mu^i,\lambda_{a\xi}) = \mu(a\xi)^{-i} \cdot g(\mu^i,\lambda)$. Hence if we denote the sum $g(\mu^i,\lambda)$ by g_i and the root of unity $\mu(x^{-1}a\xi)$ by γ, (2.7) becomes

$$(2.11) \qquad \nu(x;a) = \frac{q^{k-1}}{N} + \frac{1}{qN}\sum_{i=1}^{N-1} G_i\overline{g_i}\gamma^i, \qquad a \neq 0.$$

We saw above that μ^i acts trivially on F_e^* if and only if $i \equiv 0 \pmod{N_2}$; hence (G1) and (G3) yield the estimate

$$(2.12) \qquad \left|\nu(x;a) - \frac{q^{k-1}}{N}\right| \leq \frac{(N_1-1)+(N-N_1)q^{\frac{1}{2}}}{N} q^{k/2-1}.$$

Because of the relative simplicity of the formula (2.9), for the rest of the paper we shall restrict our attention to the case $a = 0$; that is, we shall content ourselves with a study of the weights of the codewords $c(x)$. (It is probable, however, that the formula (2.11) can be used to extend our results to the case $a \neq 0$ as well.) The heart of the formula (2.9) is the sum

$$\sum_{i=1}^{N_1-1} G_{N_2 i}\beta^{-N_2 ij}.$$

But $G_{N_2 i} = G(\mu^{N_2 i},\lambda)$ is just a Gauss sum for a character $\mu^{N_2 i}$ of order N_1, and β^{-N_2} is an N_1-st root of unity. Thus we change our notation and for the remainder of the paper let β denote an N_1-st root of unity, and μ a character of F_{ek}^* such that $\mu(\psi) = \beta$. We denote the Gauss sum $G(\mu^i,\lambda)$ by G_i. Our goal is thus the calculation of the sum

$$(2.13) \qquad \sum_{i=1}^{N_1-1} G_i\beta^{-ij}.$$

We shall succeed in evaluating the sum (2.13) if $p^l \equiv -1 \pmod{N_1}$ for some l; if N_1 is prime and p generates the quadratic residues mod N_1; if $N_1 = 2$; $N_1 = 3$; or $N_1 = 4$.

3. THE SEMIPRIMITIVE CASE

If $N_1 > 2$ and if there exists an integer l such that $p^l \equiv -1 \pmod{N_1}$ *) we say that p is *semiprimitive* mod N_1. In this case it is possible to determine the Gauss sums of order N_1 in the field F_{2l} explicitly. The result is (G6) that for $i=1,2,\ldots,N_1-1$,

$$G(\mu^i,\lambda) = (-1)^i p^l, \quad (N_1 \text{ even and } \frac{p^l+1}{N_1} \text{ odd}),$$
$$G(\mu^i,\lambda) = \quad p^l, \quad (N_1 \text{ odd or } \frac{p^l+1}{N_1} \text{ even}) .$$

We are however interested in the Gauss sums of order N_1 in the field F_{ek}. Since $N_1 > 2$ it follows that $2l$ is a divisor of ek, say $2lm = ek$. Thus by the theorem of DAVENPORT & HASSE (G5),

$$(3.1) \qquad \begin{cases} G_i = (-1)^{m+1+im} q^{k/2}, & (N_1 \text{ even and } \frac{p^l+1}{N_1} \text{ odd}), \\ G_i = (-1)^{m+1} q^{k/2}, & (N_1 \text{ odd or } \frac{p^l+1}{N_1} \text{ even}) . \end{cases}$$

From (3.1) it is a simple matter to compute the sums (2.13), and thus also to compute $\nu(x;0)$ from (2.9). The result is

$$(3.2) \qquad \nu(x;0) = \begin{cases} \frac{q^{k-1}-1}{N} + \frac{(-1)^{m+1}(q-1)(N_1-1)}{N} q^{k/2-1} \\ \frac{q^{k-1}-1}{N} + \frac{(-1)^{m}(q-1)}{N} q^{k/2-1} . \end{cases}$$

The first alternative holds when $j \equiv 0 \pmod{N_1}$, unless N_1 is even, $(p^l+1)/N_1$ is odd, and m is odd, in which case it holds when $j \equiv N_1/2 \pmod{N_1}$. The second alternative holds in all other cases. Incidentally it is not necessary to know m in order to use the formulas (3.2) since if the wrong

---

*) We assume that l is in fact the least positive integer with this property.

sign is used the resulting numbers will not be integers.

As an example, consider the (91,6) irreducible cyclic code over GF(3). Then $N = 8$ and $N_1 = 4$. Since $3 \equiv -1 \pmod 4$ the results of this section apply. From (3.2) it thus follows that

$$\nu(x;0) = 37 \quad \text{if} \quad \text{ind}(x) \equiv 2 \pmod 4,$$

$$\nu(x;0) = 28 \quad \text{if} \quad \text{ind}(x) \not\equiv 2 \pmod 4 .$$

4. THE QUADRATIC RESIDUE CASE

Suppose that N_1 is an odd prime and that p generates the quadratic residues mod N_1, i.e., $p^{(N_1-1)/2}$ is the least power of p congruent to 1 (mod N_1). If $N_1 \equiv 1 \pmod 4$, -1 is a quadratic residue mod N_1 and the results of section 3 apply. If $N_1 \equiv 3 \pmod 4$ the calculation of a Gauss sum of order N_1 in $F_{(N_1-1)/2}$ is not so easy, but BAUMERT & MYKKELTVEIT [2] have shown (cf. G10) that $G(\mu,\lambda) = p^{(N_1-1)/4} e^{\pm i\theta}$, where θ is as described in the appendix. We are interested in the Gauss sums of order N_1 in the field F_{ek}, however. Since $p^{ek} \equiv 1 \pmod{N_1}$ it follows that ek is divisible by $(N_1-1)/2$; we denote the quotient $2ek/(N_1-1)$ by m. Then by the Davenport-Hasse result, by choosing the primitive root ψ of F_{ek} properly,

$$(4.1) \qquad G_1 = (-1)^{m+1} q^{k/2} e^{-im\theta} = -q^{k/2} e^{im\alpha}, \qquad \alpha = \pi - \theta.$$

By (G4), $G_i = G_1$ if i is a quadratic residue mod N_1 and $G_i = \overline{G_1}$ if i is a non-residue. Hence if we denote the sum $\sum\{\beta^i : (\frac{i}{p}) = +1\}$ by η, the sum (2.13) assumes one of the three values $(N_1-1)(G_1+\overline{G}_1)$, $(G_1\eta+\overline{G}_1\overline{\eta})$, $(G_1\overline{\eta}+\overline{G}_1\eta)$. GAUSS [5] article 356 showed that η is a solution to the quadratic equation $x^2+x+(N_1+1)/4 = 0$, i.e., $\eta = (-1\pm\sqrt{-N})/2$. The ambiguity in η is immaterial for our purposes since the mapping $\eta \to \overline{\eta}$ merely interchanges $G_1\eta+\overline{G}_1\overline{\eta}$ and $G_1\overline{\eta}+\overline{G}_1\eta$. Hence without loss of generality we may take

$$(4.2) \qquad \eta = -\frac{(N_1+1)^{\frac{1}{2}}}{2} e^{i\rho}, \quad \tan\rho = N_1^{\frac{1}{2}}, \qquad 0 < \rho < \pi/2 .$$

It follows from (4.1) and (4.2) that the sum (2.13) assumes one of the three values $-(N_1-1)q^{k/2}\cos m\alpha$, $(N_1+1)^{\frac{1}{2}} q^{k/2} \cos(m\alpha\pm\rho)$. Thus by (2.9)

$$(4.3)\qquad \nu(x;0) = \begin{cases} \dfrac{q^{k-1}-1}{N} - \dfrac{(q-1)(N_1-1)}{N}\, q^{k/2-1} \cos m\alpha \ , \\ \dfrac{q^{k-1}-1}{N} + \dfrac{(q-1)(N_1+1)}{N}\, q^{k/2-1} \cos(m\alpha \pm \rho) \ . \end{cases}$$

The first alternative in (4.3) holds when $j \equiv \mathrm{ind}(x) \equiv 0 \pmod{N_1}$. If $\mathrm{ind}(x) \not\equiv 0 \pmod{N_1}$ the quadratic character of $\mathrm{ind}(x)$ determines whether the second or third alternative holds. Finally, the angles α and ρ are given by

$$(4.3') \qquad \begin{cases} \tan\rho = N_1^{\frac{1}{2}},\ 0 < \rho < \pi/2 \ , \\ \alpha = \pi - \theta,\ \tan\theta = bN_1^{\frac{1}{2}}/a,\ 0 < \theta < \pi \ , \\ a^2 + N_1 b^2 = 4p^{s-2t},\ a \equiv -2p^{s-t} \pmod{N_1} \ , \\ s = (N_1-1)/2,\ t = w_p(n_1)/(p-1),\ n_1 = (p^s-1)/N_1 \ . \end{cases}$$

As an example consider the (71,5) code over GF(5). Here $N = 44$ and $N_1 = 11$. Since $11 \equiv 3 \pmod 4$ and $5^6 \equiv 1 \pmod{11}$ the results of this section apply. From (4.3') we compute $\rho = \tan^{-1}\sqrt{11} = 73.221345^\circ$, $\alpha = \pi - \tan^{-1}\sqrt{11}/3 = 132.130415^\circ$. From (4.3) it then follows that

$$\nu(x;0) = \begin{cases} 21 & \mathrm{ind}(x) \equiv 0 \pmod{11} \ , \\ 16 & \mathrm{ind}(x) \equiv 1,3,4,5,9 \pmod{11} \ , \\ 11 & \mathrm{ind}(x) \equiv 2,6,7,8,10 \pmod{11} \ , \end{cases}$$

provided the primitive root ψ of $GF(5^5)$ is properly chosen. (Otherwise the second two values would be interchanged.)

5. $N_1 = 2$

We now suppose that p is odd and that $N_1 = 2$. Then since $(q^k-1)/(q-1) = 1+q+\ldots+q^{k-1} \equiv k \pmod 2$, it follows that k must be even. In the field F_1, the Gauss sum of order 2 is given by (cf. G7)

$$G(\mu,\lambda) = \begin{cases} \sqrt{p} & \text{if } p \equiv 1 \pmod 4 , \\ \sqrt{-p} & \text{if } p \equiv 3 \pmod 4 . \end{cases}$$

Hence by the Davenport-Hasse theorem, in the field F_{ek}

$$G_1 = \begin{cases} q^{k/2} & \text{if } p \equiv 3 \pmod 4,\ ek \equiv 2 \pmod 4 , \\ -q^{k/2} & \text{otherwise} . \end{cases} \tag{5.1}$$

Thus by (2.9),

$$\nu(x;0) = \frac{q^{k-1}-1}{N} \pm \frac{(q-1)}{N} q^{k/2-1} .$$

If $p \equiv 3 \pmod 4$ and $ek \equiv 2 \pmod 4$, the + sign applies when $j \equiv \mathrm{ind}(x) \equiv 0 \pmod 2$. In all other cases the + sign applies when ind(x) is odd.

As an example consider the (410,4) code over GF(9). Then $N = 16$, $N_1 = 2$. Thus by (5.2)

$$\nu(x;0) = \begin{cases} 41 & \mathrm{ind}(x) \equiv 0 \pmod 2 , \\ 50 & \mathrm{ind}(x) \equiv 1 \pmod 2 . \end{cases}$$

6. $N_1 = 3$

If $N_1 = 3$ and $p \equiv 2 \pmod 3$ the results of section 3 will apply. Thus suppose $p \equiv 1 \pmod 3$. Then $(q^k-1)/(q-1) \equiv 1+q+\ldots+q^{k-1} \equiv k \pmod 3$. Hence $k \equiv 0 \pmod 3$. According to (G8), in the field F_1, the Gauss sum of order 3 is given by

$$G(\mu,\lambda) = p^{\frac{1}{2}} e^{\pm i\theta},$$

where θ is as given in the appendix. Thus if we denote ek by m, the cubic Gauss sum G_1 in the field F_{ek} can be assumed to be

$$G_1 = (-1)^{m+1} q^{k/2} e^{-im\theta} = -q^{k/2} e^{im\alpha}, \qquad \alpha = \pi-\theta. \tag{6.1}$$

With reasoning similar to that of section 4, it follows that the sum (2.13) assumes one of the three values $-2q^{k/2}\cos m\alpha$, $-2q^{k/2}\cos(m\alpha\pm 2\pi/3)$. Hence

$$(6.2)\qquad \nu(x;0) = \begin{cases} \dfrac{q^{k-1}-1}{N} - \dfrac{2(q-1)}{N} q^{k/2-1} \cos m\alpha\ , \\ \dfrac{q^{k-1}-1}{N} - \dfrac{2(q-1)}{N} q^{k/2-1} \cos(m\alpha\pm 2\pi/3)\ , \end{cases}$$

where α is determined by

$$(6.2')\qquad \begin{cases} \alpha = \pi-\theta,\ \tan 3\theta = 3b\sqrt{3}/a,\ 0 < \theta < \pi/3 \\ a^2+27b^2 = 4p,\ a \equiv 1 \pmod 3\ . \end{cases}$$

As an example consider the (57,3) code over GF(7). Then $N = 6$, $N_1 = 3$. From (6.2') we calculate $a = 1$, $b = \pm 1$, $\theta = 26.368868^\circ$, $\alpha = 153.631132^\circ$. It follows from (6.2) that

$$\nu(x;0) = \begin{cases} 9 & \text{ind}(x) \equiv 0 \pmod 3\ , \\ 12 & \text{ind}(x) \equiv 1 \pmod 3\ , \\ 3 & \text{ind}(x) \equiv 2 \pmod 3\ , \end{cases}$$

provided the primitive root ψ of $GF(7^3)$ has been chosen properly. If the wrong primitive root is selected the second and third values of $\nu(x;0)$ would be interchanged.

7. $N_1 = 4$

If $N_1 = 4$ and $p \equiv 3 \pmod 4$ the results of section 3 apply. Hence we assume that $p \equiv 1 \pmod 4$. Also, since $(q^k-1)/(q-1) = 1+q+\ldots+q^{k-1} \equiv k \pmod 4$, it follows that $k \equiv 0 \pmod 4$. By (G9) the Gauss sum of order 4 in F_1 is $p^{\frac{1}{2}} e^{\pm i\theta}$, where θ is given in the appendix. Hence in the field F_{ek}

$$(7.1)\qquad G_1 = -q^{k/2} e^{im\theta}\ .$$

By (5.1), $G_2 = -q^{k/2}$. Furthermore $G_3 = \overline{G_1}$, and so the sum (2.13) is

$$q^{k/2}((-1)^{j+1} - 2\cos(m\theta - j\pi/2)).$$

Thus $\nu(x;0)$ is given by

$$(7.2)\qquad \nu(x;0) = \begin{cases} \dfrac{q^{k-1}-1}{N} - \dfrac{(q-1)(1\pm 2\cos m\theta)}{N} q^{k/2-1}, & j\equiv 0,2 \pmod 4. \\ \dfrac{q^{k-1}-1}{N} + \dfrac{(q-1)(1\pm 2\sin m\theta)}{N} q^{k/2-1}, & j\equiv 1,3 \pmod 4. \end{cases}$$

The angle θ is determined by

$$(7.2')\qquad \begin{cases} \tan 4\theta = 4ab/(4b^2-a^2),\ 0 < \theta < \pi/4\ , \\ a^2+4b^2 = p,\ a \equiv 1 \pmod 4\ . \end{cases}$$

As an example consider the (39,4) code over GF(5). Then $N = 16$, $N_1 = 4$. The angle $\theta = \frac{1}{4}\tan^{-1}\frac{4}{3} = 13.282526^\circ$. Thus the values of $\nu(x;0)$ are

$$\nu(x;0) = \begin{cases} 5 & j \equiv 0 \pmod 4\ , \\ 11 & j \equiv 1 \pmod 4\ , \\ 8 & j \equiv 2 \pmod 4\ , \\ 7 & j \equiv 3 \pmod 4\ . \end{cases}$$

APPENDIX: SOME PROPERTIES OF GAUSS SUMS

Let $F_k = GF(p^k)$ and let N be an integer dividing p^k-1. If ψ is a primitive root of F_k and if $x \neq 0$ is an element of F_k, we define the index of x (with respect to ψ) as $ind(x) = i$, where $\psi^i = x$ and $i \in \{0,1,\ldots,p^k-2\}$. Let ζ be a complex p-th root of unity, i.e., $\zeta = \exp(2\pi ih/p)$ for some $h \in \{0,1,\ldots,p-1\}$. Then for any $b \in F_k$ we may define a character of the additive group $F_k^{(+)}$ of F_k by

$$(A1)\qquad \lambda(x) = \zeta^{T_1^k(bx)}.$$

Similarly if β is any complex N-th root of unity we define a character of the multiplicative group $F_k^{(\cdot)}$ of non-zero elements of F_k by

$$(A2)\qquad \mu(x) = \beta^{ind(x)}.$$

It turns out that every character of $F_k^{(+)}$ has the form (A1), and every character of $F_k^{(\cdot)}$ of order N has the form (A2).

The *Gauss sum* of the characters μ,λ in F_k is now defined by

$$(A3) \qquad G(\mu,\lambda) = \sum_{\substack{x\in F_k \\ x\neq 0}} \mu(x)\ \lambda(x)\ .$$

This is also called a Gauss sum of order N. Such sums (sometimes they are called *Lagrange resolvents*) have been intensively studied since GAUSS considered the special case N = 2, f = 1 in [5, article 356], and much is known about them. The article by IWASAWA [8] is a very good introduction to the subject, but is difficult to obtain. It contains an especially good treatment of Stickleberger's theorem. The recent book by IRELAND & ROSEN [7] is also a good source of information. The last chapter of HASSE's book [6] is not as elementary as the other two accounts but is more complete. In this appendix we shall list, but not prove, the results needed for this paper.

The first result deals with the sums $G(\mu,\lambda)$ where either μ or λ is trivial.

$$(G1) \qquad G(\mu,\lambda) = \begin{cases} p^k-1 & \text{if } \mu = 1,\ \lambda = 1, \\ 0 & \text{if } \mu \neq 1,\ \lambda = 1, \\ -1 & \text{if } \mu = 1,\ \lambda \neq 1. \end{cases}$$

Property (G1) is quite easy to prove from the definition (A3).

The next property shows how changing the character λ affects the value of $G(\mu,\lambda)$. For this purpose we denote the character in (A1) by λ_b.

$$(G2) \qquad G(\mu,\lambda_b) = \mu(b)^{-1} G(\mu,\lambda_1) \quad \text{if } b \neq 0.$$

Property (G2) also follows easily from the definition.

The next property is the first really interesting property of Gauss sums, but it does not lie very deep. A proof may be found on p.92 of LANG [9].

$$(G3) \qquad \begin{cases} G(\mu,\lambda)G(\mu^{-1},\lambda) = \mu(-1)\ p^k, \quad \text{and so} \\ \\ |G(\mu,\lambda)| = p^{k/2}, \text{ if } \mu \neq 1,\ \lambda \neq 1. \end{cases}$$

The next property shows that certain automorphisms of $Q(\beta)$ leave $G(\mu,\lambda)$ invariant.

(G4) $\qquad G(\mu^{p^i},\lambda) = G(\mu,\lambda) \qquad$ for all $i=0,1,2,\ldots$,

Property (G4) follows directly from (A3) and the fact that $T_1^e(x^{p^i}) = T_1^e(x)$.

We now come to the remarkable theorem of DAVENPORT & HASSE [4], the result which is easily the most important for the applications to irreducible cyclic codes. Let k' be the order of $p \pmod N$, i.e., the smallest positive integer such that $p^{k'} \equiv 1 \pmod N$. Since also $p^k \equiv 1 \pmod N$, k is divisible by k', say $k = k'm$, and so $F_{k'}$ is a subfield of F_k. If λ' is a non-trivial character on $F_{k'}^{(+)}$ and μ' is a character on $F_{k'}^{(\cdot)}$ of order N, we may form the sum $G(\mu',\lambda')$ in $F_{k'}$. We now "lift" the characters μ',λ' from $F_{k'}$ to F_k by defining

$$\lambda(x) = \lambda'(T_{k'}^k(x))$$

$$\mu(x) = \mu'(N_{k'}^k(x)) .$$

The character λ is non-trivial on F_k since the trace $T_{k'}^k$ is onto, and the character μ is of order N since the norm $N_{k'}^k$ is onto. The theorem of Davenport and Hasse shows that there is a simple relationship between the sum $G(\lambda,\mu)$ in the field F_k and the sum $G(\mu',\lambda')$ in the smaller field $F_{k'}$:

(G5) $\qquad G(\mu,\lambda) = (-1)^{m+1} G(\mu',\lambda')^m .$

Since every character of order N in F_k can be obtained by lifting a character of order N from $F_{k'}$, and since the value of $G(\mu,\lambda)$ is not materially dependent upon which character $\lambda \neq 1$ is chosen, (G5) allows us to compute *any* Gauss sum of order N in F_k in terms of a Gauss sum in a smaller field. A proof of the Davenport-Hasse theorem was given by MCELIECE & RUMSEY in [11].

The remaining results concern the explicit calculation of certain Gauss sums. The first is what BAUMERT & MCELIECE [1] called the *semiprimitive case*.

Here it is assumed that there exists an integer l such that $p^l \equiv -1 \pmod N$. It is further assumed that β in (A2) is a primitive N-th root of unity. Then STICKLEBERGER [13, §3.6 and 3.10] showed that for any $\lambda \neq 1$, in F_{2l}

$$\text{(G6)}\qquad \begin{cases} G(\mu^i,\lambda) = (-1)^i p^l, & (N_1 \text{ even and } \frac{p^l+1}{N_1} \text{ odd}), \\ G(\mu^i,\lambda) = p^l, & (N_1 \text{ odd or } \frac{p^l+1}{N_1} \text{ even}). \end{cases}$$

This result also appears as a lemma (p.168) in BAUMERT & MCELIECE [1].

The remaining results of this section concern the explicit determination of the sum $G(\mu,\lambda)$ in a variety of other special cases. However, the determination is not as explicit as (G6) in general for the following reasons. First, we did not specify exactly either the N-th root of unity β or the primitive root ψ of F_k. This uncertainty will in general cause an ambiguity in the determination of $G(\mu,\lambda)$ of an automorphism of the field $Q(\beta)$. Second, we did not specify either the p-th root of unity ζ or the choice of b in the definition (A1) of the character λ. Property (G2) shows that this uncertainty will in general cause an ambiguity of a multiplicative factor of an N-th root of unity. The first ambiguity is inevitable because there is no "canonical" way to choose a primitive root of a finite field. However, the second ambiguity is in principle resolvable if we take $\zeta = \exp(2\pi i/p)$, $b = 1$. Unfortunately even if this is done the problem of resolving the ambiguity is in general intractable. GAUSS spent a year on the case $N = 2$, and the case $N = 3$ has never been resolved (CASSELS [3]). Thus we shall not specify the p-th root of unity ζ exactly, and accept this nagging but essentially harmless ambiguity.

We now come to the earliest result about Gauss sums. It is due to GAUSS himself [5, art.356]. We assume p is odd, $N = 2$, $f = 1$. In this case

$$\text{(G7)}\qquad G(\mu,\lambda) = \begin{cases} \pm\sqrt{p}, & p \equiv 1 \pmod 4, \\ \pm\sqrt{-p}, & \text{otherwise.} \end{cases}$$

Incidentally, GAUSS succeeded in determining the doubtful sign in (G7) as +, if $\zeta = \exp(2\pi i/p)$.

Next we assume that $p \equiv 1 \pmod 3$ and that μ is a non-trivial character of order 3 in F_1. Then with a suitable choice of ζ,

$$\text{(G8)}\qquad \begin{cases} G(\mu,\lambda) = \sqrt{p}\, e^{\pm i\theta}, \\ \tan 3\theta = 3b\sqrt{3}/a,\ 0 < \theta < \pi/3, \\ a^2+27b^2 = 4p,\ a \equiv 1 \pmod 3,\ b > 0. \end{cases}$$

These results appear implicity in GAUSS [5, art.358] but a clearer proof is given by HASSE [6, §20.4]. It turns out that the representation $4p = a^2+27b^2$ in integers a,b is unique except for the signs of a and b. The sign of a is determined by the congruence $a \equiv 1 \pmod 3$. The sign of b cannot be determined, since it reflects the uncertainty in the choice of ψ.

Next we consider primes $\equiv 1 \pmod 4$ and consider Gauss sums of order 4 in F_1. Here the result is that with a suitable choice of ζ,

$$\text{(G9)} \qquad \begin{cases} G(\mu,\lambda) = \sqrt{p}\, e^{\pm i\theta}, \\ \tan 4\theta = 4ab/(4b^2-a^2),\ 0 < \theta < \pi/4, \\ a^2+4b^2 = p,\ a \equiv 1 \pmod 4,\ b > 0. \end{cases}$$

Once again this result is essentially due to GAUSS, see HASSE [6, §20.4].

We now come to a result about Gauss sums which is apparently new, and which has arisen from the study of irreducible cyclic codes. The assumption is that N is an odd prime $\equiv 3 \pmod 4$ and that p generates the quadratic residues mod N, i.e., that the order of p (mod N) is $s = (N-1)/2$. In this case BAUMERT & MYKKELTVEIT [2] have shown that in the field F_s

$$\text{(G10)} \qquad \begin{cases} G(\mu,\lambda) = p^{s/2}\, e^{\pm i\theta}, \\ \tan\theta = b\sqrt{N}/a,\ 0 < \theta < \pi, \\ a^2+Nb^2 = 4p^{s-2t},\ a \equiv -2p^{s-t} \pmod N. \end{cases}$$

In (G10) the integer t is determined as follows. Let $n = (p^s-1)/N$, and let $n = \sum_{j=0}^{s-1} n_j p^j$ be the expansion of n in the base p. Then t is the p-weight of n divided by p-1, i.e.,

$$\text{(G10')} \qquad (p-1)t = w_p(n) = n_0+n_1+\ldots+n_{s-1}.$$

Alternatively BAUMERT & MYKKELTVEIT show that if $r_1, r_2, \ldots, r_s$ are the quadratic residues mod N reduced mod N, i.e., $0 < r_1 < r_2 < \ldots < r_s < N$, then

$$\text{(G10")} \qquad Nt = r_1+r_2+\ldots+r_s.$$

The key to the proof of (G10) is the determination of the highest power of

p which devides $G(\mu,\lambda)$. It turns out that a famous theorem of STICKLEBERGER [13, §6] shows that this highest power is $\min(w_p(n), w_p((N-1)n))$. The fact that $w_p(n) < w_p((N-1)n)$ follows easily from the famous theorem of Gauss that for primes N of the form 4k+3 there are more quadratic residues in the range (0,N/2) than in the range (-N/2,0). (For a proof of this result see WEYL [14,ch.IV].)

LIST OF SYMBOLS

p, a prime

q, a power p^e of p

n, an integer not divisible by p

k, the least positive integer such that $q^k \equiv 1 \pmod n$

$F^{(+)}$, the additive group of the field F

$F^{(\cdot)}$, the multiplicative group of non-zero elements of F

F^*, the set of non-zero elements of F

F_i, the field $GF(p^i)$

T_i^{ij}, the trace of F_{ij} over F_i

N_i^{ij}, the norm of F_{ij} over F_i

$N = (q^k-1)/n$

$N_1 = g.c.d.(N, 1+q+\ldots+q^{k-1})$

$N_2 = N/N_1$

β, a complex primitive N or N_1-st root of unity

ζ, a complex primitive p-th root of unity

ξ, an element of F_{ek} with $T_e^{ek}(\xi) = 1$

ψ, a primitive root in F_{ek}

θ, a primitive n-th root of unity in F_{ek}, usually $\theta = \psi^N$

ind(x), the least positive integer j such that $\psi^j = x$ in F_{ek}

c(x), a codeword (vector) whose i-th component is $T_e^{ek}(x\theta^i)$, $x \in F_{ek}$ $(i=0,1,\ldots,n-1)$

$\nu(x;a)$, the number of components of c(x) equal to a

μ, a character of $F_{ek}^{(\cdot)}$, usually of order N or N_1

λ, a character of $F_{ek}^{(+)}$, usually $\lambda(x) = \zeta^{T_1^{ek}(x)}$

λ_b, a character defined by $\lambda_b(x) = \lambda(bx)$

$G(\mu,\lambda)$, the Gauss sum $\sum\{\mu(x)\lambda(x) : x \in F_{ek}^*\}$

G_i, the Gauss sum $G(\mu^i,\lambda)$

g_i, the Gauss sum $\sum\{\mu^i(x)\lambda(x) : x \in F_e^*\}$

REFERENCES

[1] BAUMERT, L.D. & R.J. MCELIECE, *Weights of irreducible cyclic codes*, Information and Control, 20 (1972) 158-175.

[2] BAUMERT, L.D. & J. MYKKELTVEIT, *Weight distributions of some irreducible cyclic codes*, DSN Progress Report 16 (1973) 128-131 (published by Jet Propulsion Laboratory, Pasadena, California).

[3] CASSELS, J.W.S., *On Kummer sums*, Proc. London Math. Soc. (3), 21 (1970) 19-27.

[4] DAVENPORT, H. & H. HASSE, *Die Nullstellen der Kongruenzzetafunktionen in gewissen zyklischen Fallen*, J. Reine Angew. Math., 172 (1935) 151-182.

[5] GAUSS, C.F., *Disquisitiones Arithmeticae*, English translation published by Yale University Press, New Haven, 1966.

[6] HASSE, H., *Vorlesungen über Zahlentheorie*, Springer-Verlag, Berlin, 1964.

[7] IRELAND, K. & M.I. ROSEN, *Elements of number theory*, Bogden and Quigley, Tarrytown-on-Hudson, 1972.

[8] IWASAWA, K., *Stickelberger's theorem on Gauss sums*, Notes taken by J. SMITH at the National Science Foundation Advanced Science Seminar, Bowdoin College, 1966.

[9] LANG, S., *Algebraic number theory*, Addison Wesley, Reading, 1970.

[10] LINT, J.H. VAN, *Coding theory*, Lecture Notes in Mathematics 201, Springer-Verlag, Berlin etc., 1971.

[11] MCELIECE, R.J. & H. RUMSEY, Jr., *Euler products, cyclotomy and coding*, J. Number Theory, 4 (1972) 302-311.

[12] OGANESIAN, S.S. & V.G. YAGDZIAN, *A class of optimal cyclic codes with base* p, Problemy Peredači Informacii, 8 (1972), vyp. 2, 109-111 (in Russian).

[13] STICKELBERGER, L., *Ueber eine Verallgemeinerung der Kreisteilung*, Math. Ann., 37 (1890) 321-367.

[14] WEYL, H., *Algebraic theory of numbers*, Annals of Mathematics Studies 1, Princeton University Press, Princeton, 1940.