Matrix Theory & Applications

PROCEEDINGS OF SYMPOSIA
IN APPLIED MATHEMATICS
Volume 40

MATRIX THEORY AND APPLICATIONS

Charles R. Johnson, Editor

American Mathematical Society
Providence, Rhode Island

LECTURE NOTES PREPARED FOR THE
AMERICAN MATHEMATICAL SOCIETY SHORT COURSE

MATRIX THEORY AND APPLICATIONS

HELD IN PHOENIX, ARIZONA
JANUARY 10–11, 1989

The AMS Short Course Series is sponsored by the Society's Committee on Employment and Educational Policy (CEEP). The series is under the direction of the Short Course Advisory Subcommittee of CEEP.

Library of Congress Cataloging-in-Publication Data

Matrix theory and applications/Charles R. Johnson, editor.
p. cm. – (Proceedings of symposia in applied mathematics, ISSN 0160-7634; v. 40. AMS short course lecture notes)
"Held in Phoenix, Arizona, January 10–11, 1989"–T.p. verso.
Includes bibliographical references.
ISBN 0-8218-0154-6 (alk. paper)
1. Matrices congresses. I. Johnson, Charles R. II. American Mathematical Society. III. Series: Proceedings of symposia in applied mathematics; v. 40. IV. Series: Proceedings of symposia in applied mathematics. AMS short course lecture notes.
QA188.M395 1990 90-30-584
512.9′434–dc20 CIP

1980 *Mathematics Subject Classification*. (1985 *Revision*).
Primary 15-02.

Printed in the United States of America.
The paper used in this book is acid-free and falls within the guidelines
established to ensure permanence and durability. ♾
Portions of this publication were typeset using $\mathcal{A}\mathcal{M}\mathcal{S}$-TEX,
the American Mathematical Society's TEX macro system.

10 9 8 7 6 5 4 3 2 1   95 94 93 92 91 90

AMS SHORT COURSE LECTURE NOTES
Introductory Survey Lectures
A Subseries of Proceedings of Symposia in Applied Mathematics

PROCEEDINGS OF SYMPOSIA IN APPLIED MATHEMATICS

Volume 20 **THE INFLUENCE OF COMPUTING ON MATHEMATICAL RESEARCH AND EDUCATION**
Edited by J. P LaSalle (University of Montana, August 1973)

Volume 19 **MATHEMATICAL ASPECTS OF COMPUTER SCIENCE**
Edited by J. T. Schwartz (New York City, April 1966)

Volume 18 **MAGNETO-FLUID AND PLASMA DYNAMICS**
Edited by H. Grad (New York City, April 1965)

Volume 17 **APPLICATIONS OF NONLINEAR PARTIAL DIFFERENTIAL EQUATIONS IN MATHEMATICAL PHYSICS**
Edited by R. Finn (New York City, April 1964)

Volume 16 **STOCHASTIC PROCESSES IN MATHEMATICAL PHYSICS AND ENGINEERING**
Edited by R. Bellman (New York City, April 1963)

Volume 15 **EXPERIMENTAL ARITHMETIC, HIGH SPEED COMPUTING, AND MATHEMATICS**
Edited by N. C. Metropolis, A. H. Taub, J. Todd, and C. B. Tompkins (Atlantic City and Chicago, April 1962)

Volume 14 **MATHEMATICAL PROBLEMS IN THE BIOLOGICAL SCIENCES**
Edited by R. Bellman (New York City, April 1961)

Volume 13 **HYDRODYNAMIC INSTABILITY**
Edited by R. Bellman, G. Birkhoff and C. C. Lin (New York City, April 1960)

Volume 12 **STRUCTURE OF LANGUAGE AND ITS MATHEMATICAL ASPECTS**
Edited by R. Jakobson (New York City, April 1960)

Volume 11 **NUCLEAR REACTOR THEORY**
Edited by G. Birkhoff and E. P. Wigner (New York City, April 1959)

Volume 10 **COMBINATORIAL ANALYSIS**
Edited by R. Bellman and M. Hall, Jr. (New York University, April 1957)

Volume 9 **ORBIT THEORY**
Edited by G. Birkhoff and R. E. Langer (Columbia University, April 1958)

Volume 8 **CALCULUS OF VARIATIONS AND ITS APPLICATIONS**
Edited by L. M. Graves (University of Chicago, April 1956)

Volume 7 **APPLIED PROBABILITY**
Edited by L. A. MacColl (Polytechnic Institute of Brooklyn, April 1955)

Volume 6 **NUMERICAL ANALYSIS**
Edited by J. H. Curtiss (Santa Monica City College, August 1953)

Volume 5 **WAVE MOTION AND VIBRATION THEORY**
Edited by A. E. Heins (Carnegie Institute of Technology, June 1952)

Volume 4 **FLUID DYNAMICS**
Edited by M. H. Martin (University of Maryland, June 1951)

Volume 3 **ELASTICITY**
Edited by R. V. Churchill (University of Michigan, June 1949)

Volume 2 **ELECTROMAGNETIC THEORY**
Edited by A. H. Taub (Massachusetts Institute of Technology, July 1948)

Volume 1 **NON-LINEAR PROBLEMS IN MECHANICS OF CONTINUA**
Edited by E. Reissner (Brown University, August 1947)

Preface

As a subfield of mathematics, matrix theory continues to enjoy a renaissance that has accelerated during the past decade, though its roots may be traced much further back. This is due in part to stimulation from a variety of applications and to the considerable interplay with other parts of mathematics, but also to a great increase in the number and vitality of specialists in the field. As a result, the once popular misconception that the subject has been fully researched has been largely dispelled. The interest on the part of the American Mathematical Society and the approximately 140 participants in the Short Course (at the January 1989 Phoenix Meeting) on which this volume is based is a reflection of this change. The steady growth in quality and volume of the subject's three principal journals, Linear Algebra and its Applications, Linear and Multilinear Algebra, and the SIAM Journal on Matrix Analysis and Applications is another. Approximately 500 different authors have published in one of these three journals in the last two years. Geographically, strong research centers in matrix theory have developed recently in Portugal and Spain, Israel, the Netherlands, Belgium, and Hong Kong.

The purpose of the Short Course was to present a sample of the ways in which modern matrix theory is stimulated by its interplay with other subjects. Though the course was limited to seven speakers, the "other subjects" represented included combinatorics, probability theory, statistics, operator theory and control theory, algebraic coding theory, partial differential equations, and analytic function theory. Among other important examples, numerical analysis, optimization, physics and economics are, unfortunately, at most lightly touched. There is no limit to the specific examples that might be cited.

One of the ingredients in the recent vitality of matrix theory is the variety of points of view and tools brought to the subject by researchers in different areas. This is responsible for a number of important trends in current research. For example, the notion of majorization (mentioned in the talk by Olkin) has become pervasive in a historically brief period of time. The trend away from the "basis-free" point of view is illustrated by work in combinatorial matrix theory (Brualdi, Johnson), the Hadamard product

(Horn) and nonnegative matrices (Mukherjea). There are many quite worthy issues that are at least excruciatingly difficult to view in a basis-free way, and freedom from the basis-free view has opened many exciting avenues of research. On the other hand, recognition of the "right" problem dependant symmetries can provide vital insight (Diaconis). The synergy between matrix theory and systems theory has had a tremendous impact on both, and on the now highly mathematically driven field of electrical engineering (Gohberg). The immense variety of tools and problems illustrates a reason for use of the term "matrix theory" or "matrix analysis" in place of "linear algebra". A large portion of current work is neither primarily linear nor primarily algebraic in nature. No point of view on what the subject is or where it is going could, or should, be without substantial disagreement. This only reflects the remarkable breadth of interest enjoyed by the subject. For an historical perspective on the nature and role of the subject the reader might enjoy the prefaces to each of the following: *Recent Advances in Matrix Theory* (Schneider 1964); *A Survey of Matrix Theory and Matrix Inequalities* (Marcus and Minc, 1964); *Linear Algebra and its Applications*, volume 1 (Alan Hoffman, 1968); and *Matrix Analysis* (Horn and Johnson, 1985). A glimpse of the contagious appeal of the subject is communicated by Olga Taussky in her November 1988 *Monthly* article "How I Became a Torchbearer for Matrix Theory".

As organizer, I would like to again thank each of the speakers for a contribution that will advance both the subject and the general understanding of it. The significant time necessary to prepare both a talk and then subsequent paper is much appreciated. The community would also like to thank the American Mathematical Society for recognizing, and providing a forum for, the subject.

Charles R. Johnson
College of William and Mary

Table of Contents

Proceedings of Symposia in Applied Mathematics
Volume 40, 1990

THE MANY FACETS OF COMBINATORIAL MATRIX THEORY

Richard A. Brualdi

ABSTRACT. I take a very broad view of combinatorial matrix theory. Combinatorial matrix theory is concerned with the use of matrix theory and linear algebra in proving combinatorial theorems and in describing and classifying combinatorial constructions, and it is also concerned with the use of combinatorial ideas and reasoning in the finer analysis of matrices and with intrinsic combinatorial properties of matrix arrays.

1. THE DETERMINANT. It should come as no surprise that combinatorial theory and matrix theory could interact to form a subject called "combinatorial matrix theory". One of the first matrix functions one studies is the determinant, and the determinant has a very combinatorial definition: The determinant of an n by n matrix

(1.1) $$A = (a_{ij} : i = 1,...,n;\ j = 1,...,n)$$

is given by

(1.2) $$\det A = \sum_{\pi} (\operatorname{sign} \pi) \prod_{i=1}^{n} a_{i\pi(i)}$$

where the summation is over all permutations π of $\{1,...,n\}$, and where $\operatorname{sign} \pi$ is $+1$ or -1 according as π is an even or odd permutation. We may impart even more combinatorial spirit to the determinant through the use of directed graphs (or digraphs).

A digraph Γ of order n has a set of n vertices, usually taken to be the set $\{1,...,n\}$, and a set $\mathcal{A}$ of arcs where each arc is an ordered pair (i,j) of not necessarily distinct vertices. The arcs of Γ can be regarded as a subset of the positions of the matrix (1.1). The entry a_{ij} at the position (i,j) of A is the weight assigned by A to the arc (i,j) of Γ. If $a_{ij} = 1$ or 0 according as (i,j) is or is not an arc of Γ $(i,j = 1,...,n)$, then A is the adjacency matrix of Γ. We may use these arc weights in order to assign weights to other objects associated with Γ. However this may be done, we define the weight of a set of objects to be the sum of the weights of the objects in the set.

This paper was written during a period in which the author was partially supported by National Science Foundation Grant No. DMS–8521521.

Let k be a positive integer. A cycle γ of length k is a sequence

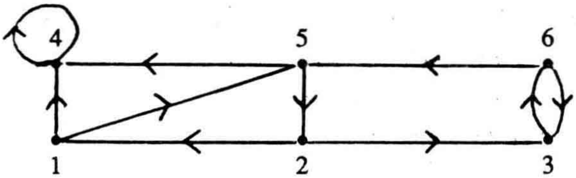
(1.3) $$(i_1,\ldots,i_k, i_1)$$

of vertices such that $i_1,\ldots,i_k$ are distinct and $(i_1,i_2),\ldots,(i_{k-1},i_k),(i_k,i_1)$ are arcs of Γ (referred to as the arcs of γ). (If $i_1,\ldots,i_k$ are not assumed to be distinct, then (1.3) is called a circuit.) We define the weight of a cycle (1.3) to be

(1.4) $$-a_{i_1i_2} \cdots a_{i_{k-1}i_k} a_{i_ki_1},$$

the negative of the products of the weights of its arcs. When we refer to disjoint cycles, we mean cycles with no vertex in common. The weight of a pairwise disjoint union of cycles is the product of the weights of the individual cycles.

(1.5) Example: Let Γ be the digraph of order 6 pictured below. Then



(1.6) $$\{(1,5,2,1), (3,6,3), (4,4)\}$$

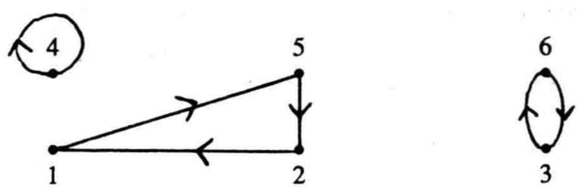
is a pairwise disjoint union of three cycles of Γ having weight

$$(-a_{15}a_{52}a_{21})(-a_{36}a_{63})(-a_{44}) = -a_{15}a_{21}a_{36}a_{44}a_{52}a_{63} \cdot$$ □

To a permutation π of $\{1,\ldots,n\}$ we associate a permutation digraph $\Gamma(\pi)$ of order n whose set of arcs is $\{(i,\pi(i)): i = 1,\ldots,n\}$. There is a one–to–one correspondence between the cycles of the permutation π and the cycles of the digraph $\Gamma(\pi)$. For example, let π be the permutation of $\{1,\ldots,6\}$ defined by

$$\pi(1) = 5, \pi(2) = 1, \pi(3) = 6, \pi(4) = 4, \pi(5) = 2, \pi(6) = 3.$$

Then $\Gamma(\pi)$ is pictured as



with the three cycles (1.6). Moreover

$$\pi = (1\ 5\ 2)(4)(3\ 6)$$

is a representation of π as a product of pairwise disjoint permutation cycles of lengths 3,1, and 2 respectively.

Each permutation digraph $\Gamma(\pi)$ is the pairwise disjoint union of cycles encompassing all vertices and its weight is

$$(1.7)\qquad (-1)^{\#\pi} \prod_{i=1}^{n} a_{i\pi(i)}$$

where $\#\pi$ is the <u>number of cycles</u> of π. Suppose the cycles of π have lengths $\ell_1,\ldots,\ell_k$, respectively, where $k = \#\pi$. Then

$$\text{sign } \pi = (-1)^{(\ell_1-1)+\cdots+(\ell_k-1)} = (-1)^{n-k} = (-1)^n(-1)^k .$$

Hence the weight (1.7) of $\Gamma(\pi)$ equals

$$(1.8)\qquad (\text{sign } \pi) \prod_{i=1}^{n} (-a_{i\pi(i)}) .$$

Let $\mathcal{P}(n)$ denote the set of all permutation digraphs of order n. From the above discussion we conclude that

$$(1.9)\qquad \det(-A) = \text{weight}(\mathcal{P}(n)) ,$$

a very combinatorial interpretation of the determinant.

Let I_n denote the identity matrix of order n. Then $\det(I_n-A)$ is the sum of the determinants of all principal submatrices of $-A$. It now follows that

$$(1.10)\qquad \det(I_n-A) = \text{weight}(\mathcal{P}^*(n))$$

where $\mathcal{P}^*(n)$ denotes the set of all permutation digraphs whose vertices form a subset of $\{1,\ldots,n\}$.

While the above combinatorial interpretation of the determinant has been known for some time, we have relied on the particular description in [Ze].

It was Jurkat and Ryser [JuRy1] who first showed that the determinant of an n by n matrix can be represented as the product of n matrices. Since the determinant is a scalar, the first matrix in the product must have one row and the last matrix must have one column. A very short and very combinatorial derivation of this factorization was given in [BrSh] within the general framework of a graded poset. We consider here only the graded poset P of all subsets of $\{1,\ldots,n\}$ partially ordered by inclusion and graded by cardinality.

Let P_i denote the collection of all i–element subsets of $\{1,\ldots,n\}$ $(i = 0,1,\ldots,n)$, and let $\alpha_1^{(i)}, \alpha_2^{(i)}, \ldots, \alpha_{\binom{n}{i}}^{(i)}$ be a listing, say in lexicographic order, of the sets in P_i. The chains of length n (the maximum possible length) in P are of the form

$$(1.11)\qquad \phi = \alpha_{i_0=1}^{(0)} \subset \alpha_{i_1}^{(1)} \subset \alpha_{i_2}^{(2)} \subset \cdots \subset \alpha_{i_n=1}^{(n)} = \{1,\ldots,n\} .$$

Let j_k be the unique element of $\alpha^{(k)}_{i_k}$ not belonging to $\alpha^{(k-1)}_{i_{k-1}}$ $(k = 1,\ldots,n)$. Then

$$j_1 j_2,\ldots,j_n \tag{1.12}$$

is a permutation of $\{1,\ldots,n\}$. Conversely, given a permutation (1.12) we obtain a chain (1.11) by defining $\alpha^{(0)}_{i_0} = \phi$ and $\alpha^{(k)}_{i_k} = \{j_1,\ldots,j_k\}$ $(k = 1,\ldots,n)$.

We use the n by n matrix (1.1) to assign weights to pairs consisting of an $(i-1)$–element set $\alpha^{(i-1)}_k$ in P_{i-1} and an i–element set $\alpha^{(i)}_\ell$ in P_i:

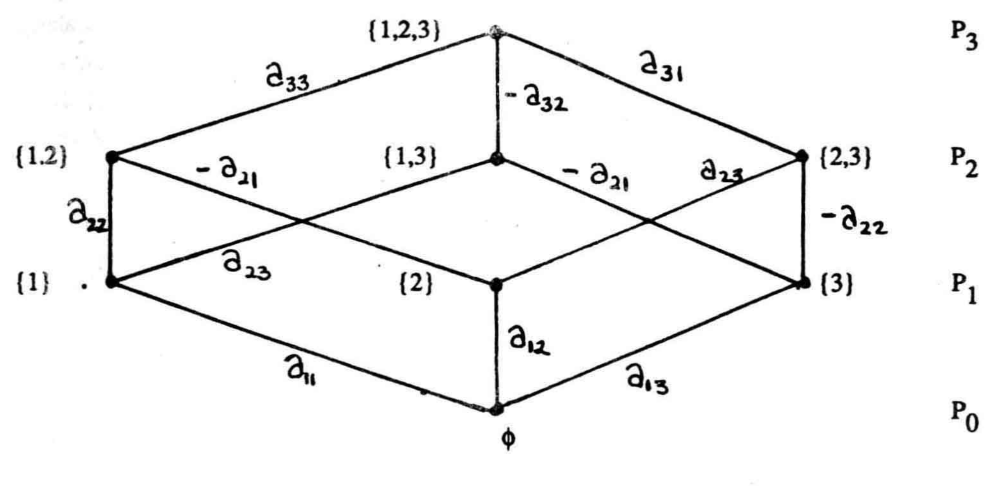
$$w^{(i)}_{k\ell} = \begin{cases} 0, & \text{if } \alpha^{(i-1)}_k \not\subset \alpha^{(i)}_\ell \\ (-1)^{c_j} a_{ij}, & \text{if } \alpha^{(i-1)}_k \cup \{j\} = \alpha^{(i)}_\ell \\ & \text{and } c_j \text{ elements of } \alpha^{(i-1)}_k \\ & \text{are greater than } j. \end{cases} \tag{1.13}$$

The $\binom{n}{i-1}$ by $\binom{n}{i}$ matrix

$$W^{(i)} = (w^{(i)}_{k\ell} : k = 1,\ldots,\tbinom{n}{i-1};\ \ell = 1,\ldots,\tbinom{n}{i})$$

is the <u>weighted incidence matrix</u> for P_{i-1} and P_i $(i = 1,\ldots,n)$. We observe that the entries of $W^{(i)}$ depend only on the entries in row i of A.

(1.14) <u>Example</u>: Let $n = 3$. Then the partially ordered set P with weights assigned as above is pictured below. The (formally) zero weights do not appear in this diagram which is the usual Hasse diagram of P with labels (the weights) on its edges.



The chain $\phi \subset \{3\} \subset \{1,3\} \subset \{1,2,3\}$ corresponds to the permutation 3,1,2. □

We now define the <u>weight of the chain</u> (1.11) to be

(1.15) $$w^{(1)}_{i_0 i_1} w^{(2)}_{i_1 i_2} \cdots w^{(n)}_{i_{n-1} i_n} .$$

If (1.11) corresponds to the permutation π in (1.12) then denoting the number of inversions of π by $\mathrm{inv}(\pi)$, we see that (1.15) equals

$$(-1)^{\mathrm{inv}(\pi)} a_{1j_1} a_{2j_2} \cdots a_{nj_n} = (\mathrm{sign}\ \pi) \prod_{i=1}^{n} a_{i\pi(i)} .$$

Let $\mathcal{C}_n$ denote the set of all chains in P having length n. It now follows that

(1.16) $$\det A = \mathrm{weight}(\mathcal{C}_n) .$$

On the other hand the matrix

$$W^{(1)} W^{(2)} \cdots W^{(n)}$$

is a 1 by 1 matrix whose unique entry equals

(1.17) $$\sum_{i_{n-1}} \cdots \sum_{i_2} \sum_{i_1} w^{(1)}_{1 i_1} w^{(2)}_{i_1 i_2} \cdots w^{(n)}_{i_{n-1} 1} .$$

Since the product $w^{(1)}_{1 i_1} w^{(2)}_{i_1 i_2} \cdots w^{(n)}_{i_{n-1} 1}$ is zero if $\phi = \alpha^{(0)}_1, \alpha^{(1)}_{i_1}, \ldots, \alpha^{(n-1)}_{i_{n-1}}, \alpha^{(n)}_1 = \{1,\ldots,n\}$ is not a chain, we conclude that

(1.18) $$\mathrm{weight}(\mathcal{C}_n) = W^{(1)} W^{(2)} \cdots W^{(n)} .$$

Combining (1.16) and (1.18) we obtain

(1.19) $$\det A = W^{(1)} W^{(2)} \cdots W^{(n)} ,$$

a <u>matrix factorization of the determinant</u>.

(1.20) <u>Example</u>: Let n = 3. Then the factorization (1.19) is $W^{(1)} W^{(2)} W^{(3)}$ where

$$W^{(1)} = \phi \begin{matrix} \{1\} & \{2\} & \{3\} \\ [a_{11} & a_{12} & a_{13}] \end{matrix} ,$$

$$W^{(2)} = \begin{matrix} \\ \{1\} \\ \{2\} \\ \{3\} \end{matrix} \begin{matrix} \{1,2\} & \{1,3\} & \{2,3\} \\ \left[\begin{matrix} a_{22} \\ -a_{21} \\ 0 \end{matrix}\right. & \begin{matrix} a_{23} \\ 0 \\ -a_{21} \end{matrix} & \left.\begin{matrix} 0 \\ a_{23} \\ -a_{22} \end{matrix}\right] \end{matrix} ,$$

$$W^{(3)} = \begin{array}{c} \\ \{1,2\} \\ \{1,3\} \\ \{2,3\} \end{array} \begin{array}{c} \{1,2,3\} \\ \begin{bmatrix} a_{33} \\ -a_{32} \\ a_{31} \end{bmatrix} \end{array} \qquad \square$$

If we define the weights $w_{k\ell}^{(i)}$ by

$$w_{k\ell}^{(i)} = \begin{cases} 0, & \text{if} \quad \alpha_k^{(i-1)} \not\subset \alpha_\ell^{(i)} \\ a_{ij}, & \text{if} \quad \alpha_k^{(i-1)} \cup \{j\} = \alpha_\ell^{(i)} \end{cases}$$

then we obtain as above matrices $W^{(1)}, W^{(2)}, \ldots, W^{(n)}$ where now

$$W^{(1)}W^{(2)} \cdots W^{(n)} = \text{per } A ,$$

the <u>permanent</u> of the matrix A.

2. Matrix–theoretic Proofs of Combinatorial Theorems.

In this section we illustrate how matrix theory has been used to prove theorems in combinatorics. These theorems have formulations which appear to be "purely combinatorial", but the proofs given show that the theorems have a hidden matrix–theoretic meaning, which is provided by the incidence matrix.

Let $X = \{x_1,\ldots,x_v\}$ be a nonempty set of v elements, and let $X_1,\ldots,X_b$ be b not necessarily distinct subsets of X. The <u>incidence matrix</u> of the subsets $X_1,\ldots,X_b$ of X is the b by v (0,1)–matrix

$$A = (a_{ij} : 1 \le i \le b;\quad 1 \le j \le v)$$

where $a_{ij} = 1$ if $x_j \in X_i$ and $a_{ij} = 0$ if $x_j \notin X_i$. Row i of the incidence matrix displays the elements of the set X_i while column j displays the sets containing the element x_j. The incidence matrix gives a complete description of the subsets $X_1,\ldots,X_b$ of the set X. The importance of the incidence matrix in combinatorial theory rests with the observation that combinatorial statements concerning the subsets $X_1,\ldots,X_b$ of X can often be formulated as algebraic statements about the matrix A.

We now describe a class of configurations which are central in combinatorial theory. Let k and t be nonnegative integers and let λ be a positive integer. Then the subsets $X_1,\ldots,X_b$ of $X = \{x_1,\ldots,x_v\}$ form a <u>t–design</u> provided:

(2.1) Each X_i contains k elements;

(2.2) For each subset T of t elements of X, there are exactly λ of the sets $X_1,...,X_b$ which contain T.

A t–design is usually denoted as $S_\lambda(t,k,v)$, and the sets $X_1,...,X_v$ are called blocks. A t–design with $\lambda = 1$ is called a Steiner system. A 2–design is commonly known as a balanced incomplete block design (BIBD). Trivial examples of t–designs are obtained by taking $X_1,...,X_b$ to be the family of all the distinct k–element subsets of X (thus $b = \binom{v}{k}$ and $\lambda = \binom{v-t}{k-t}$)). The following are some basic facts about t–designs [Ha,Wil1,Wil2]

(2.3) $b = \lambda\binom{v}{t}/\binom{k}{t}$.

(2.4) A t–design $S_\lambda(t,k,v)$ is also an i–design $S_{\lambda_i}(i,k,v)$ where $\lambda_i = \lambda\binom{v-i}{t-i}/\binom{k-i}{t-i}$ $(i = 0,1,...,t)$. In particular, the λ_i are integers.

According to (2.3) the integer b is determined by v,k,λ, and t. A fundamental problem in combinatorial theory is to determine if a t–design $S_\lambda(t,k,v)$ exists for given v,k,λ and t. A 1–design $S_\lambda(1,k,v)$ is a very general combinatorial configuration consisting of b subset $X_1,...,X_b$ of a v–element set $X = \{x_1,...x_v\}$ where each X_i contains exactly k elements and each x_j is contained in exactly λ blocks. By (2.4) a t–design $S_\lambda(t,k,v)$ with $t \geq 2$ is both a 1–design $S_{\lambda_1}(1,k,v)$ and a 2–design $S_{\lambda_2}(2,k,v)$. We now consider 2–designs $S_\lambda(2,k,v)$ and denote the number λ_1 of blocks containing a specified element by r. By (2.3) and (2.4) $bk = rv$ and $r(k-1) = \lambda(v-1)$.

Let $A = (a_{ij})$ be the b by v incidence matrix of an $S_\lambda(2,k,v)$, and let A^T denote the transpose of A. The (i,j)–entry of A^TA is

$$\sum_{u=1}^{b} a_{ui}a_{uj} ,$$

which equals the number λ of blocks containing x_i and x_j if $i \neq j$ and equals the number r of blocks containing x_i if $i = j$. Let I_m denote the m by m identity matrix, and let J_{mn} denote the m by n matrix each of whose entries equals 1 (abbreviated to J_m if $m = n$). Then the properties (2.1) and (2.2) of a 2–design are entirely equivalent to the matrix equations

(2.5) $$AJ_{vb} = kJ_b,\quad A^TA = (r-\lambda)I_v + \lambda J_v.$$

Since each element of X is in exactly r blocks, A also satisfies

(2.6) $$J_{v,b}A = rJ_v.$$

We now assume that $v > k$, equivalently that $r > \lambda$. The matrix $(r-\lambda)I_v + \lambda J_v$ has eigenvalues $r+(v-1)\lambda$ and $\lambda-r$ (v–1 times), and it is easily checked that its inverse satisfies

(2.7) $$((r-\lambda)I_v + \lambda J_v)^{-1} = \frac{1}{r-\lambda}(I_v - \frac{\lambda}{rk}J_v).$$

It follows from (2.5) that the v by v matrix A^TA is nonsingular. Since A is a b by v matrix we obtain

(2.8) Fisher's inequality: Let $v > k$. Then the number b of blocks and the number v of points in a 2–design $S_\lambda(2,k,v)$ satisfy $b \geq v$. □

We recall the following basic fact from linear algebra:

(2.9) Let C be an m by n matrix with rank equal to n. Then the orthogonal projection from the m–dimensional real vector space m onto the column space W of C is given by the m by m matrix

(2.9a) $$P = C(C^TC)^{-1}C^T.$$

The matrix P is symmetric and idempotent with eigenvalues 0 (m–n times) and 1 (n times). The orthogonal projection of $\mathbb{R}^m$ onto the orthogonal complement of W is given by the m by m matrix $Q = I_m - P$. The matrix Q is symmetric and idempotent with eigenvalues 0 (n times) and 1 (m–n times). In particular, both P and Q are positive semidefinite. □

We apply (2.9) to the incidence matrix A of an $S_2(\lambda,k,v)$. The matrix A^TA is nonsingular, and using (2.5), (2.6), (2.7), and (2.9a) we calculate the orthogonal projections P and Q:

$$\begin{aligned} P &= A(A^TA)^{-1}A^T \\ &= A((r-\lambda)I_v + \lambda J_v)^{-1}A^T \\ &= A\frac{1}{r-\lambda}(I_v - \frac{\lambda}{rk}J_v)A^T; \end{aligned}$$

hence

(2.10) $$P = \frac{1}{r-\lambda}(AA^T - \frac{\lambda k}{r}J_b),$$

and

(2.11) $$Q = I_b - P = I_b - \frac{1}{r-\lambda}(AA^T - \frac{\lambda k}{r} J_b).$$

Because Q is positive semidefinite each of its principal submatrices has a nonnegative determinant. An m by m principal submatrix of Q is obtained as follows. Let $X_1',\dots,X_m'$ be m of the blocks of a 2–design $S_2(\lambda,k,v)$, and let X_i' and X_j' have μ_{ij} elements in common $(i,j = 1,\dots,m)$. Let $U = (\mu_{ij}: i,j = 1,\dots,m)$. Then

$$Q_m = I_m - \frac{1}{r-\lambda}(U - \frac{\lambda k}{r} J_m)$$

is a principal submatrix of Q, and hence

(2.12) $$\det Q_m \geq 0.$$

The 2^b-1 inequalities of the form (2.12) are Connor's inequalities [Co,Hal,Wil2]. The diagonal entries of Q_m are equal to $(r-k)/r$. For $i \neq j$ the (i,j)–entry of Q_m equals $(\lambda k - r\mu_{ij})/r(r-\lambda)$. The case $m = 1$ of (2.12) asserts that $r \geq k$; because $bk = rv$, this is equivalent to Fisher's inequality (2.8). The case $m = 2$ of (2.12) is equivalent to the statement that the number μ of elements common to two different blocks satisfies

(2.13) $$(r-k)(r-\lambda) \geq |\lambda k - r\mu|.$$ □

A symmetric 2–design is an $S_2(\lambda,k,v)$ with $b=v$ (and hence $r=k$). For a symmetric design (2.13) is equivalent to $\mu = \lambda$. Thus in a symmetric design two distinct blocks have exactly λ elements in common. The incidence matrix A of a symmetric design thus satisfies the matrix equations

(2.14) $$J_v A = AJ_v = kJ_v, \quad AA^T = A^TA = (k-\lambda)I_v + \lambda J_v.$$

The second set of equations in (2.14) implies that if there is a symmetric 2–design $S_2(\lambda,k,v)$, then the matrices I_v and $(k-\lambda)I_v + \lambda J_v$ are rationally congruent. This observation leads to the following necessary conditions of Bruck, Ryser, and Chowla [Hal,Ry2] for the existence of a symmetric $S_2(\lambda,k,v)$:

(2.15) If v is odd, then the equation

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

has a solution in integers x,y, and z not all equal to 0. □

A pairwise balanced design is obtained by removing the condition (2.1) in the definition of a 2–design. Thus subsets $X_1,\dots,X_b$ of $X = \{x_1,\dots,x_v\}$ form a pairwise balanced design of index $\lambda \geq 1$ if and only if the incidence matrix A satisfies