

NO-A177 228

NEW PROPERTIES OF ORTHOGONAL ARRAYS AND THEIR
STATISTICAL APPLICATIONS (U) ILLINOIS UNIV AT CHICAGO
CIRCLE DEPT OF MATHEMATICS STATISTIC A S HEDAYAT
OCT 86 TR-86-15 AFOSR-TR-87-0066 F/G 12/1

1/1

UNCLASSIFIED

NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AD-A177 228



THE UNIVERSITY OF ILLINOIS AT CHICAGO

AFOSR-TR-87-0066

New Properties of Orthogonal Arrays
and Their Statistical Applications

By

A.S. Hedayat*

Approved for public release;
distribution unlimited.

AIR FORCE OFFICE OF SCIENTIFIC RESEARCH (AFSC)
NOTICE OF TRANSMITTAL TO DTIC
This technical report has been reviewed and is
approved for public release IAW AFR 190-12.
Distribution is unlimited.
MATTHEW J. KERPER
Chief, Technical Information Division

DTIC FILE COPY

DEPARTMENT OF MATHEMATICS,
STATISTICS, AND COMPUTER SCIENCE

DTIC
ELECTE
FEB 24 1987
S
D

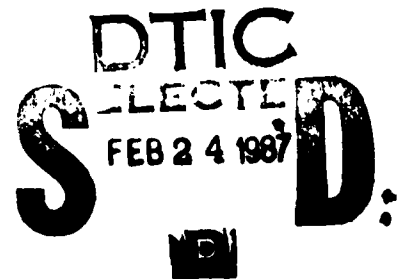
New Properties of Orthogonal Arrays
and Their Statistical Applications

By

A.S. Hedayat*

Department of Mathematics, Statistics, and Computer Science
University of Illinois at Chicago
Chicago, Illinois 60680

STATISTICAL LABORATORY TECHNICAL REPORT
No. 86-15, October, 1986



*

Research is sponsored by Grant AFOSR 85-0320.

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION UNCLASSIFIED			1b. RESTRICTIVE MARKINGS		
2a. SECURITY CLASSIFICATION AUTHORITY N/A			3. DISTRIBUTION/AVAILABILITY OF REPORT Approved for public release; Distribution unlimited.		
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A					
4. PERFORMING ORGANIZATION REPORT NUMBER(S)			5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 87-0066		
6a. NAME OF PERFORMING ORGANIZATION University of Illinois at Chicago		6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM		
6c. ADDRESS (City, State and ZIP Code) P.O. Box 4348 Chicago, IL 60680			7b. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling Air Force Base, D.C. 20332-6448		
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR		8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR 85-0320		
8c. ADDRESS (City, State and ZIP Code) Bldg. 410 Bolling Air Force Base, D.C. 20332-6448			10. SOURCE OF FUNDING NOS.		
			PROGRAM ELEMENT NO. 6.110F	PROJECT NO. 2304	TASK NO. A5
11. TITLE (Include Security Classification) New Properties of Orthogonal Arrays and Their Statistical Applications					
12. PERSONAL AUTHOR(S) A.S. Hedavat					
13a. TYPE OF REPORT Interim		13b. TIME COVERED FROM _____ TO _____		14. DATE OF REPORT (Yr., Mo., Day) October 1986	
15. PAGE COUNT 10					
16. SUPPLEMENTARY NOTATION					
17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Orthogonal arrays; fractional factorials; flexible orthogonal arrays; line graphs.		
FIELD	GROUP	SUB. GR.			
19. ABSTRACT (Continue on reverse if necessary and identify by block number) It is shown that an orthogonal array of strength t is more than a fractional array design of resolution $t + 1$. The practical usefulness of this result is shown. The notion of <i>flexible orthogonal arrays</i> of strength t is introduced and its practical usefulness is demonstrated. An efficient way of generating the design and information matrices associated with orthogonal arrays in the context of orthogonal polynomial models is presented.					
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐			21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED		
22a. NAME OF RESPONSIBLE INDIVIDUAL Major Brian W. Woodruff			22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027		22c. OFFICE SYMBOL NM

New Properties of Orthogonal Arrays and Their Statistical Applications

A. HEDAYAT*

ABSTRACT

It is shown that an orthogonal array of strength t is more than a fractional factorial design of resolution $t + 1$. The practical usefulness of this result is shown. The notion of *flexible orthogonal arrays* of strength t is introduced and its practical usefulness is demonstrated. An efficient way of generating the design and information matrices associated with orthogonal arrays in the context of orthogonal polynomial models is presented.



Accession For	
NTIS CRA&I	☒
DTIC TAB	☐
Unannounced	☐
Justification	
By	
Distribution /	
Availability Codes	
Dist	Avail and/or Special
A-1	

Key words: Orthogonal arrays; fractional factorials; flexible orthogonal arrays; line graphs.

AMS Subject Classification Codes: Primary 62K15, Secondary 05B15.

*Research is sponsored by Grants AFOSR 85-0320.

An invited talk to be presented in the First International Conference on Statistical Computing to be held March 30 - April 2, 1987 in Cesme, Izmir, Turkey.

1. INTRODUCTION

Fractional factorial designs are among the most utilized statistical designs by practitioners. In a series of three papers Rao (1946, 1947, 1949) identified special types of fractions with a great deal of symmetry and many desirable statistical properties. These special fractions are now known as orthogonal arrays. Numerous beautiful results have been obtained by both statisticians and mathematicians on this topic. These results are published in a wide variety of journals and presented in many different styles. In their forthcoming research monograph Hedayat and Stufken (1986) have presented the various results on this fascinating subject in a unified and comprehensive way.

This paper is divided into four sections. In section 2 we will review basic definitions and terminology of the subject. In section 3 we will present an efficient way of preparing design and information matrices associated with orthogonal arrays under orthogonal polynomial models. Section 4 contains a new result and a new concept. The new result states that with orthogonal arrays of strength t we can orthogonally estimate other parametric vectors besides the one which is advocated in the literature. We have also identified some types of orthogonal arrays which are somewhere between orthogonal arrays of strength t and $t + 1$. We call them *flexible orthogonal arrays* of strength t . Practical applications of such arrays are also pointed out. These results are very useful in preparing line graphs of Taguchi for orthogonal arrays.

2. PRELIMINARIES

We begin by giving some of the basic definitions and terminology. Let S be a set of s symbols, coded by $0, 1, \dots, s - 1$. A $k \times N$ array A with entries from S is called an orthogonal array, denoted by $OA(N, k, s, t)$, if each $t \times N$ subarray of A has the property that every possible $t \times 1$ vector with entries from S appears equally often (say λ times) in the columns of the chosen subarray. The integer λ is called the index of the array, while N, k, s and t are said to be the parameters of the array. If $\lambda = 1$ the array is said to be of index unity. The relation $N = \lambda s^t$ is an immediate consequence of the definition.

The reader may verify that the following array is an $OA(8, 4, 2, 3)$, an orthogonal array of index unity.

0	1	1	1	1	0	0	0
1	0	1	1	0	1	0	0
1	1	0	1	0	0	1	0
1	1	1	0	0	0	0	1

The terminology that is used in the literature for the parameters of an orthogonal array is as follows: N is called the size of the array or the number of runs, assemmbles or treatment combinations; k is called the number of constraints or factors in the arrays; s is called the number of levels or symbols; t is called the strength of the array.

3. ORTHOGONAL ARRAYS UNDER ORTHOGONAL POLYNOMIAL MODELS

As a fractional factorial designs orthogonal arrays are highly efficient if orthogonal polynomial models are postulated for the responses under study. It was indeed in this context that Rao introduced the concept of orthogonal arrays in the literature. The purpose of this section is two-fold. To present the orthogonal polynomial models associated with orthogonal arrays so that our result in section 4 will be easily understood by the readers of this paper. We also present an efficient way, recently discovered by Hedayat and Stufken (1986), for generating the design and information matrices associated with orthogonal arrays. This new method is very useful in analyzing data obtained under orthogonal arrays and in studying the optimality of such arrays.

In the remainder of this section we will have to assume that the reader has some familiarity with the standard concepts in factorial experiments and the theory of ANOVA. For a quick review of these concepts a useful reference is Raktoe, Hedayat and Federer (1981). Assume that our interest is in an experiment based on k controllable factors, each of which can be set at s different levels. The levels are coded by $0, 1, \dots, s - 1$. A treatment combination is an experimental condition to which each factor contributes a level. For example $(i_1, \dots, i_j, \dots, i_k)$ is a treatment combination with the j^{th} factor at level i_j . Thus there are s^k distinct treatment combinations. These can be exhibited column wise in a $k \times s^k$ array, which we will denote by ρ and refer to as a minimal complete factorial based on k factors at s levels each.

Let g be a treatment combination and y_g the observed response under the experimental condition g . We will assume that y_g is a random variable,

$$y_g = f'(g)\theta + \varepsilon_g,$$

where $f'(g)$ is a row vector of real known functions of g and θ is a column vector of unknown parameters. The random error component ε_g will be assumed to have mean zero and unknown variance σ^2 .

The structure of $f'(g)$ is an important consideration in practice. If all factors are quantitative factors a popular way of modeling y_g is to structure $f'(g)$ via the orthogonal polynomial model (see for example Chapter 4 of Raktue, Hedayat and Federer (1981)). We will use this model from here on. If we perform the experiment using all s^k treatment combinations, the resulting vector of observations, Y_ρ , can be expressed as

$$(3.1) \quad Y_\rho = X_\rho \beta_\rho + \varepsilon_\rho,$$

where X_ρ is the design matrix and β_ρ is the vector of general mean, main effects and interactions. The entries in β_ρ , following the common notation, are

$$\phi_1^{i_1} \phi_2^{i_2} \dots \phi_k^{i_k}, \quad i_j \in \{0, 1, \dots, s-1\}.$$

We will now present an easy way to find the entries of X_ρ under the orthogonal polynomial model. This model is especially useful if we are only interested in a fraction of the minimal complete factorial design. Let $x_0, x_1, \dots, x_{s-1}$ be the actual levels of the first factor. Form the following matrix

$$X_1 = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^{s-1} \\ 1 & x_1 & x_1^2 & \dots & x_1^{s-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{s-1} & x_{s-1}^2 & \dots & x_{s-1}^{s-1} \end{bmatrix}.$$

Orthogonalize the columns of X_1 from left to right and call the resulting matrix M_1 . Similarly obtain $M_2, \dots, M_{k-1}$ corresponding to the other $k-1$ factors. Denote the entry in the i^{th} row and j^{th} column of M_ℓ by $n_{ij}^{(\ell)}$, $0 \leq i, j \leq s-1$, $1 \leq \ell \leq k$. Then the entry in X_ρ corresponding to the treatment combination $(x_{i_1}, x_{i_2}, \dots, x_{i_k})$ and parameter $\phi_1^{j_1} \phi_2^{j_2} \dots \phi_k^{j_k}$ is given by

$$(3.2) \quad n_{i_1 j_1}^{(1)} \cdot n_{i_2 j_2}^{(2)} \dots n_{i_k j_k}^{(k)}.$$

If the s levels of the factors are equally spaced, the orthogonalized matrices M can be obtained from the available literature. Table 4.1 in Raktoc, Hedayat and Federer (1981) lists such matrices for $2 \leq s \leq 7$, which is sufficient for most practical purposes. The same table can be used for $s = 8$ or 9 if a polynomial model of degree less than or equal to 5 is fitted for each factor. We demonstrate some of the above ideas in the following example.

Example 3.1. Let $s = 3, k = 2$. Under equally spaced levels, coded by $0, 1, 2$, we have

$$\rho = \begin{matrix} & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 \\ 0 & 0 & 1 & 2 & 0 & 1 & 2 & 0 & 1 & 2 \end{matrix}$$

$$M_1 = M_2 = \begin{matrix} & 0 & 1 & 2 \\ 0 & \begin{pmatrix} 1 & -1 & 1 \\ 1 & 0 & -2 \\ 1 & 1 & 1 \end{pmatrix} \\ 1 & \\ 2 & \end{matrix}$$

$$\beta'_\rho = (\phi_1^0 \phi_2^0, \phi_1^0 \phi_2^1, \phi_1^0 \phi_2^2, \phi_1^1 \phi_2^0, \phi_1^1 \phi_2^1, \phi_1^1 \phi_2^2, \phi_1^2 \phi_2^0, \phi_1^2 \phi_2^1, \phi_1^2 \phi_2^2).$$

Thus for example, the entry in X_ρ corresponding to treatment combination $(1, 2)$ and parameter $\phi_1^2 \phi_2^0$ is, from (3.2) equal to

$$n_{12}^{(1)} n_{20}^{(2)} = (-2) \cdot (1) = -2.$$

Completing the example we obtain

$$\begin{bmatrix} Y_{00} \\ Y_{01} \\ Y_{02} \\ Y_{10} \\ Y_{11} \\ Y_{12} \\ Y_{20} \\ Y_{21} \\ Y_{22} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & 0 & -2 & -1 & 0 & 2 & 1 & 0 & -2 \\ 1 & 1 & 1 & -1 & -1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 0 & 0 & 0 & -2 & 2 & -2 \\ 1 & 0 & -2 & 0 & 0 & 0 & -2 & 0 & 4 \\ 1 & 1 & 1 & 0 & 0 & 0 & -2 & -2 & -2 \\ 1 & -1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 0 & -2 & 1 & 0 & -2 & 1 & 0 & -2 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \beta_\rho + \epsilon_\rho.$$

4. NEW PROPERTIES OF ORTHOGONAL ARRAYS

It is often impractical or impossible due to other restrictions (money, time, etc.) to use all possible treatment combinations in an experiment. Clearly we would like to select a fraction of the s^k treatment combinations, where we will allow that some of the selected combinations are used more than once in the experiment. If A is an $OA(N, k, s, t)$ based on $S = \{0, 1, \dots, s-1\}$, then the columns of A form such a fraction of the minimal complete factorial.

Under the orthogonal polynomial model (3.1), possibly neglecting some of the parameters in β_ρ , we can write

$$(4.1) \quad Y_A = X_A \beta_A + \epsilon_A.$$

Here β_A is the vector of parameters obtained from β_ρ by deleting those parameters that can be neglected, while X_A is the design matrix obtained as in (3.2), but now only using the treatment combinations in A and the vector of parameters β_A .

Example 4.1. Let A be the following $OA(8, 4, 2, 3)$:

0	1	1	1	1	0	0	0
1	0	1	1	0	1	0	0
1	1	0	1	0	0	1	0
1	1	1	0	0	0	0	1

Let β_A consists of the general mean, the main effects and the two factor interactions between factors 1 and 2, 1 and 3, and 1 and 4. In our notation this means

$$\beta_A = (\phi_1^0 \phi_2^0 \phi_3^0 \phi_4^0, \phi_1^1 \phi_2^0 \phi_3^0 \phi_4^0, \phi_1^0 \phi_2^1 \phi_3^0 \phi_4^0, \phi_1^0 \phi_2^0 \phi_3^1 \phi_4^0, \\ \phi_1^0 \phi_2^0 \phi_3^0 \phi_4^1, \phi_1^1 \phi_2^1 \phi_3^0 \phi_4^0, \phi_1^1 \phi_2^0 \phi_3^1 \phi_4^0, \phi_1^1 \phi_2^0 \phi_3^0 \phi_4^1).$$

Since $M_1 = M_2 = M_3 = M_4 = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ we obtain by (3.2):

$$\begin{bmatrix} Y_{0111} \\ Y_{1011} \\ Y_{1101} \\ Y_{1110} \\ Y_{1000} \\ Y_{0100} \\ Y_{0010} \\ Y_{0001} \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & 1 & -1 \\ 1 & 1 & -1 & -1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & 1 & -1 & 1 & -1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{bmatrix} \beta_A + \epsilon_A$$

It is interesting to observe that in the above example $X'_A X_A = 8I$, implying that the general mean, the main effects and the 3 two factor interactions can be estimated orthogonally, assuming that other parameters can be neglected. This property is not based on coincidence. We will soon see a result that explains this property. In the meantime the reader is invited to verify that the same property holds if we replace the interaction between factors 1 and 4 in β_A by the interaction between factors 2 and 3.

We will proceed by giving some of the statistical properties of orthogonal arrays as fractional factorial designs under the orthogonal polynomial model, but for a concise formulation we need one more concept. If r is even, then a fractional factorial design is said to be of resolution r if all factorial effects involving $(r-2)/2$ or fewer factors are estimable, assuming that all factorial effects involving $(r+2)/2$ or more factors can be neglected. If r is odd, the design is said to be of resolution r if all effects involving $(r-1)/2$ or fewer factors are estimable, assuming that the remaining effects can be neglected. The usual convention about the general mean is that for odd r it is among the parameters to be estimated, while for even r it is not among those to be estimated, nor among those that can be neglected. So for example a resolution III design is one for which the general mean and main effects can be estimated, assuming that all other effects can be neglected. A resolution IV design allows us to estimate all main effects under the presence of the general mean and all two factor interactions, while all other effects are neglected.

The main reason that orthogonal arrays are interesting in the above described statistical context is due to the following well known result.

THEOREM 4.1. *An orthogonal array of strength t is a design of resolution $t+1$. Moreover the concerned effects are orthogonally estimable, while the general mean can always be included in the effects to be estimated, both for even and odd t .*

A proof of this result can, for example, be found in Raktoe, Hedayat and Federer (1981) (Theorem 13.1). This is a very nice and powerful result. It tells us for example that in an $OA(8,4,2,3)$ the general mean and main effects can be estimated orthogonally, assuming that three and four factor interactions are absent. However a property as in Example 3.2 cannot be explained by this result.

For this purpose we formulate here a more general result. Let x be a k -dimensional $(0,1)$ -vector, i.e., $x' = (x_1, x_2, \dots, x_k)$ with $x_i \in \{0,1\}$. With factorial effects corresponding to x we will mean those effects $\phi_1^{i_1} \phi_2^{i_2} \dots \phi_k^{i_k}$ for which $i_j = 0$ if and only if $x_j = 0, 1 \leq j \leq k$.

Now assume that B and C are two disjoint sets of k -dimensional $(0,1)$ -vectors (C may be the empty set) such that

1. If $x, y \in B$, then $|\{i \in \{1, \dots, k\} : x_i = 1 \text{ or } y_i = 1\}| \leq t$.
2. If $x \in B, y \in C$, then $|\{i \in \{1, \dots, k\} : x_i = 1 \text{ or } y_i = 1\}| \leq t$.

Then we have the following theorem.

THEOREM 4.2. *If A is an orthogonal array of strength t , all effects corresponding to vectors in B can be estimated orthogonally assuming that effects corresponding to vectors in C are the only other ones that are not neglected. The zero vector, corresponding to the general mean, can always be chosen in B .*

It is easy to see that this is a generalization of Theorem 4.1. Moreover if we choose $B = \{(0,0,0,0)', (1,0,0,0)', (0,1,0,0)', (0,0,1,0)', (0,0,0,1)', (1,1,0,0)', (1,0,1,0)', (1,0,0,1)'\}$ and $C = \emptyset$ we obtain an explanation for Example 3.2.

More information on the structure of the orthogonal array A can even lead to more general conclusions. Instead of a general formulation we just illustrate this by an example.

Example 4.3. Let A be the following orthogonal array

Factor								
1	0	0	0	0	1	1	1	1
2	0	0	1	1	0	0	1	1
3	0	1	0	1	0	1	0	1
4	0	1	0	1	1	0	1	0
5	0	1	1	0	1	0	0	1

This is an $OA(8,5,2,2)$. We claim that the following 8 effects can be estimated orthogonally, assuming that no other effects are present: the general mean, the main effects, and the two factor interactions between the factors 1 and 2 and 1 and 5. We can not conclude this from our previous result; indeed we do not claim that our statement is valid for any $OA(8,5,2,2)$, only for this particular one. An explanation for the validity of our statement is as follows: select any two effects to be estimated and look at the rows in A corresponding to the factors that are involved. These rows form each time one or more copies of a minimal complete factorial. For example, if we select the main effect of 2 and the interaction between 1 and 5, we see that the rows corresponding to the factors 1, 2 and 5 form one copy of a minimal complete factorial. Similar for other effects. Due to this property the validity of our statements follows. We call orthogonal arrays with such additional features as *flexible orthogonal arrays*. The study of the existence and construction of flexible orthogonal arrays is currently under investigation.

Closing Remarks. Our Theorem 4.2 and the notion of flexible orthogonal arrays are useful in preparing and efficiently cataloging line graphs associated with orthogonal arrays. The usefulness of such graphs for practical applications of orthogonal arrays is nicely demonstrated by Taguchi and Wu (1979).

REFERENCES

- HEDAYAT, A. AND STUFKEN, J., (1986). Orthogonal Arrays and Their Statistical Applications. (A monograph to be published.)
- RAKTOE, B.L., HEDAYAT, A. AND FEDERER, W.T., (1981). Factorial Designs. John Wiley and Sons, New York.
- RAO, C.R., (1946). Hypercubes of strength d leading to confounded designs in factorial experiments. *Bull. Calc. Math. Soc.*, **38**, 67-78.
- RAO, C.R., (1947). Factorial experiments derivable from combinatorial arrangements of arrays. *J. Roy. Statist. Soc.*, **9**, 128-139.
- RAO, C.R., (1949). On a class of arrangements. *Proc. Edinburgh Math. Soc.*, **8**, 119-125.
- TAGUCHI, G. AND WU, Y., (1979). Introduction to Off-Line Quality Control. Central Japan Quality Control Association.

END

3-87

DTIC