JOHN ARTHUR RAFTER

Ph. D.

CONTRIBUTIONS TO THE THEORY AND CONSTRUCTION OF PARTIALLY BALANCED ARRAYS

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

JOHN ARTHUR RAFTER

1971

This is to certify that the

thesis entitled

CONTRIBUTIONS TO THE THEORY AND
CONSTRUCTION OF PARTIALLY BALANCED ARRAYS

presented by

John Arthur Rafter

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Statistics and
Probability

Esther Seiden
Major professor

Date August 3, 1971

ABSTRACT

CONTRIBUTIONS TO THE THEORY AND
CONSTRUCTION OF PARTIALLY BALANCED ARRAYS

By

John Arthur Rafter

A partially balanced array (PBA) (m,N,s,t) of s levels, m constraints, and strength t with index set $\Lambda_{s,t} = \{\lambda(x_1,\ldots,x_t) | x_i \in \{0,1,\ldots,s-1\},\ i = 1,\ldots,t\}$ is an $(m \times N)$ matrix with entires from a set of s elements such that in every $(t \times N)$ submatrix each of the possible s^t distinct $(t \times 1)$ vectors, $(x_1,\ldots,x_t)'$, occurs as a column $\lambda(x_1,\ldots,x_t)$ times.

The PBA can serve as a design for a fractional factorial experiment with m factors each occurring at s levels, when the effect of N treatments is under investigation. If the PBA is of strength $t = 2u$, all interactions involving u or fewer factors are estimable, assuming there is no interaction of more than u factors. If $t = 2u + 1$, all interactions involving u or fewer factors can be estimated even if interactions of $u + 1$ factors are present. In addition, the PBA is a "balanced" design in the sense that the resulting variance-covariance matrix of the estimators is invariant under a permutation of the factor symbols.

In Chapter I, the analysis of a PBA is given for the special case of $s = t = 2$. The analysis of the general PBA is a straightforward generalization.

One problem of interest for PBA's is to determine the maximum possible number of constraints for a given N, s, t and $\Lambda_{s,t}$. In Chapter II, it is shown that, for any PBA, $m \leq N$. In the event that $s = t = 2$, better bounds are derived when $\mu_1^2 > \mu_0\mu_2$, $\mu_1^2 = \mu_0\mu_2$, and $\mu_1 = 1$. Also, a general result is given which is useful in finding bounds in other cases. In each instance, it is shown that the bounds are attainable. Next, an iterative bound on the maximum number of constraints of a PBA of strength t is given. This depends on the maximum number of constraints for a PBA of strength $t-1$. Finally, it is shown that the bounds obtained for a PBA in two symbols can be useful for arrays in more than two symbols. Moreover, the bounds are shown to be attainable in certain cases.

The first result of Chapter III is a simple set of necessary and sufficient conditions for the existence of a PBA $(t+1,N,2,t)$. These conditions are employed to give a method of construction of a PBA $(m,2N,2,t+1)$ using a PBA $(m,N,2,t)$ when $t = 2u$. Furthermore, when $t = 2$, conditions are given under which an $m+1^{st}$ row can be added to the constructed array, and a necessary and sufficient condition for $m+1$ to be the maximum possible number of rows is given.

The final result of Chapter III is the construction of a PBA $(m,N,s,2)$ from a PBA $(v,b,2,2)$ with index set $\{b-2r+1, r-1, 1\}$, where $m = r$, $N = b-r$, $s = \frac{rv}{b}$, $\lambda(0,0) = b-r-(s-1)(2r-s-1)$, $\lambda(0,i) = r-s$ $(i = 1,\ldots,s-1)$, and $\lambda(i,j) = 1$ $(i,j = 1,\ldots,s-1)$. The array $(m,N,s,2)$ is shown to have the maximum possible number of constraints.

Chapter IV is devoted to the relation of PBA's to other areas of mathematics. In the first section, the existence of certain PBA's is shown to be equivalent to certain Tactical Configurations. In the second section, conditions under which a PBA will give a strongly regular graph are investigated. In the final section, Hadamard matrices are used to construct PBA's.

CONTRIBUTIONS TO THE THEORY AND CONSTRUCTION OF PARTIALLY BALANCED ARRAYS

By

John Arthur Rafter

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Statistics and Probability

1971

TO CAROLYN AND JOHNNY

ACKNOWLEDGMENTS

I wish to express my gratitude to Dr. Esther Seiden for her guidance and encouragement in the preparation of this manuscript. Her willingness to discuss any problem at any time is deeply appreciated.

I would also like to express my gratitude to Noralee Barnes whose skillful typing has been of immeasurable help throughout the research for and preparation of this thesis. Also, I wish to thank Dr. Dennis Gilliland for his help in the final review of the manuscript.

Finally, I wish to express my gratitude to the National Science Foundation and to the Department of Statistics and Probability, Michigan State University for their financial support during my stay at Michigan State University.

TABLE OF CONTENTS

CHAPTER I

ANALYSIS

1.1 INTRODUCTION

The basic principles of the subject of Design of Experiments were first presented by R.A. Fisher (see Fisher (1960)). Since that time a great number of researchers in many different disciplines have contributed to the development of the subject.

Designs are called two dimensional or multi-dimensional according as they control variation in one or more than one direction. Balanced incomplete block designs introduced by Yates (1936) and partially balanced incomplete block designs studied for the first time by Bose and Nair (1939) are two of the many designs of the first kind. Latin square, Youden square, and lattice square designs are a few of the second kind.

In this work, we will be concerned with the existence and properties of a class of designs of the first kind called partially balanced arrays. Moreover, since once a design has been constructed it must be analyzed, we will indicate in this first chapter how one might perform the analysis of a partially balanced array of strength 2.

1.2 FACTORIAL DESIGNS

In experimental design, the variables an experimenter controls are commonly called FACTORS. The number of forms or categories of a factor appearing in an experiment is called the number of LEVELS of that factor. A particular combination with one level from each factor is a TREATMENT. If all possible treatments, or a definite portion of them, is of interest, the experiment is called a FACTORIAL experiment.

A factor can be QUANTITATIVE, such as different temperatures or different doses of a drug, or it can be QUALITATIVE, such as different methods of testing or different chemical solutions. There is usually no natural order established among the different levels of a qualitative factor, but the different levels of a quantitative factor correspond to well-defined values of some numerical quantity.

In a factorial design, if all possible treatments are included, the experiment is called a COMPLETE FACTORIAL. If repeated measurements or observations are made for each treatment, it is called FACTORIAL WITH REPLICATES. If a complete factorial design has the same number of replicates, every level of a factor or every combination of levels of any given number of factors appears the same number of times, and the design is said to be BALANCED. Furthermore, the levels of one factor occur with each of the levels of any other factor with equal frequency, and the design is said to be ORTHOGONAL.

Suppose a characteristic under study is affected by several factors $F_1, F_2, \ldots, F_m$, where each factor can assume two or more different levels. For example, let F_i assume s_i levels $i = 1, \ldots, m$. Clearly there are $s_1 \cdot s_2 \cdot \ldots \cdot s_m$ possible treatments. In case all factors assume the same number of levels, s, (i.e. $s = s_i$, $i = 1, \ldots, m$), we call the design SYMMETRIC. In this case, a complete factorial design will consist of all possible s^m m-tuples of s elements. It will be called a COMPLETE s^m DESIGN.

The effect of a treatment in a factorial design is in general regarded as the sum of an over-all mean, μ, the effects of the m factors, and the effects of interactions of all orders among these factors. For example, in a 2^m factorial design, the 2^m treatments provide independent minimum variance estimates of the general mean and of the $2^m - 1$ effects:

m	main effects
$\frac{m(m-1)}{2}$	2-factor interaction effects
$\frac{m(m-1)(m-2)}{2 \cdot 3}$	3-factor interaction effects
$\vdots$	
$\frac{m(m-1)(m-2)\ldots(m-h+1)}{h!}$	h-factor interaction effects

and a single m factor interaction effect.

In a complete factorial design the required number of measurements is often beyond the resources of the investigator; or it is not feasible to carry out; or it gives more precision in the estimates of the main effects than necessary; or estimates of higher-order interaction effects are of less interest. For example, in the above 2^m design if $m = 8$, each main effect is an average over 128 combinations of the other factors. These considerations have given rise to the use of confounding and fractional replication of complete factorial designs.

The general theory of confounding in s^m factorial designs was derived by Bose and Kishen (1940) and Bose (1947). This was done by putting the s^m factorial level-combinations (treatments) into

1-to-1 correspondence with the s^m points of the m-dimensional finite Euclidean geometry EG(m,s), based on the finite field GF(s). The s levels were taken to be in 1-to-1 correspondence with the elements of GF(s). Using these correspondences and various other properties and features of EG(m,s) and the associated finite projective geometry PG(m-1,s), the results were obtained.

This work was continued by Bose and others to include the theory of fractionally replicated designs of the type s^{m-k}, obtained by taking the s^{m-k} level-combinations satisfying a set of k appropriately chosen linear equations over GF(s). The equations were chosen so as to obtain a fractionally replicated design with the property that all the n-factor and lower order interactions are estimable, assuming that the remaining higher order effects are zero. Such a design is said to be of RESOLUTION 2n+1 (see Box and Hunter (1961)). A design to estimate the n-factor and lower order effects, assuming that (n+1)-factor interactions may be non-zero and interactions of higher order are zero, is said to be of resolution 2n + 2.

Consider a $\frac{1}{s^k}$ replication of a complete s^m factorial design, where the s^{m-k} treatments chosen satisfy the above mentioned linear equations. These treatments can be represented by a $m \times s^{m-k}$ matrix. Furthermore, it can be shown that this matrix is an orthogonal array of strength t.[1]

An orthogonal array (m,N,s,t) of s levels, m constraints, strength t, and index λ is a $m \times N$ matrix with entries from a

[1] It is not true that every orthogonal array can be obtained as a solution of a set of linear equations of the kind mentioned above.

set S of s elements, $s \geq 2$, such that each $t \times N$ submatrix contains all possible $t \times 1$ column vectors of S each repeated λ times. Orthogonal arrays were first defined and studied by Rao (1947, 1950). Their importance, which was suggested above, derives from the fact that a necessary and sufficient condition that a fraction be of resolution $(t + 1)$, where the estimates of various parameters are mutually uncorrelated, is that it be an orthogonal array of strength t.

Unfortunately, although orthogonal arrays lead to a reduction in the number of treatments necessary to estimate a given set of factors, this decrease is often not large enough, with the result that they become uneconomic to use. Thus if we insist that the fractional factorial design be such that the estimates are mutually uncorrelated, then, in general, the number of treatments will be much larger than the number of effects to be estimated, and experimental costs will rise. The obvious remedy to the situation is to drop the requirement that the estimates are mutually uncorrelated.

Let L be the vector of parameters and $\hat{L}$ its estimate, and let V denote $\mathrm{Var}\ (\hat{L})$, the variance-covariance matrix of the estimates. Then if the estimates are mutually uncorrelated, V is a diagonal matrix. In view of the economic conditions discussed above, V must in general be taken to be non-diagonal. We can, however, restrict our attention to a certain class of patterned matrices and still reduce the number of treatments considerably. We shall call a matrix V "balanced" if it is a member of this class.

In a fractional factorial design, the variance-covariance matrix of the estimators is called "balanced" if it is invariant under a permutation of the factor symbols. For example, let $L' = (\mu; F_1,\ldots,F_m)$, then V is "balanced" if $\mathrm{Var}\,(\hat{F}_i)$, $\mathrm{Cov}\,(\hat{\mu},\hat{F}_i)$ and $\mathrm{Cov}\,(\hat{F}_i,\hat{F}_j)$ are independent of the indices i and j, $i \neq j$; $i,j = 1,\ldots,m$. A fractional factorial design will be called "balanced" if its variance-covariance matrix is "balanced". Srivastava (1970) has indicated that a necessary and sufficient condition for a fractional factorial design of resolution $t + 1$ to be "balanced" is that it be a partially balanced array (PBA) of strength t.

1.3 PARTIALLY BALANCED ARRAYS

A partially balanced array[1] (PBA) A with parameters $(m,N,2,t)$ and index set $(\mu_0,\mu_1,\ldots,\mu_t)$ is an $m \times N$ matrix with elements 0 and 1 (say) such that in every $t \times N$ submatrix every vector containing i nonzero elements occurs μ_i times as a column. (To obtain a "balanced" fractional factorial design from A, simply take its columns as treatments to be included in the design.) It is clear that if $\mu_i = \lambda$, $i = 0,\ldots,t$, then A is an orthogonal array of strength t. Thus a PBA is seen to be a generalization of an orthogonal array.

Partially Balanced Arrays were first defined and studied by I.M. Chakravarti (1956, 1961, 1963). In addition, a substantial

---

[1] In view of the definition of "balanced" given above, these arrays might better be named Balanced Arrays. Indeed, Chopra and Srivastava have begun doing just that (see for example Srivastava and Chopra (1971a)). On the other hand, if by balanced one refers to the fact that every combination of levels of any given number of factors appears the same number of times, then there is a logical reason for calling them partially balanced arrays, since any combination of t or fewer factors appears the same number of times. In this paper we will refer to the arrays as Partially Balanced Arrays or even less formally as PBA's.

amount of work has been done in this area by D.V. Chopra and J.N. Srivastava. We refer the interested reader to the Bibliography.

The subject of arrays in general and PBA's in particular covers a rather wide sector of present combinatorial theory, with applications in areas like the construction of statistical experimental design, the theory of error-correcting and error-detecting codes, tactical configurations, and graph theory.

A good deal of work has already been done in special branches of the general area of PBA's. For example, explicit studies on orthogonal arrays of strength 2 and 3 have been made by Bose and Bush (1952), and on strength 4 by Seiden and Zemach (1966). Special problems have been studied under other titles as well, such as mutually orthogonal Latin Squares and Hadamard matrices.

Another important special case of PBA's is the much studied area of balanced incomplete block (BIB) designs. The incidence matrix of a BIB design with parameters (v,b,r,k,λ) is identical with a PBA in two symbols, v rows, b columns, and of strength 2, where $\mu_0 = b - 2r + \lambda$, $\mu_1 = r - \lambda$ and $\mu_2 = \lambda$. Thus, every BIB design corresponds to a PBA, and conversely every PBA in two symbols and of strength 2 corresponds to a BIB design (with possibly unequal block size).

When PBA's are considered as fractional factorial designs, they are preferable to orthogonal arrays. As mentioned above, orthogonal arrays involve an undesirably large number of treatments (columns). For example, an orthogonal array of strength two, six symbols and four rows would require at least 72 columns, but for the same situation, Chakarvarti (1961) has constructed a PBA with 42 columns.

1.4 ANALYSIS OF A PBA OF STRENGTH 2

The analysis given here is presented in a somewhat different form and in more generality by Bose and Srivastava (1964).

Consider a complete 2^m factorial experiment. Let F_i $i = 1,\dots,m$ represent the i^{th} factor and f_i represent one of the two levels at which F_i can occur; for purposes of clarity this level will be called the second level. We will signify the first level by absence of the corresponding letter. Thus the treatment $f_1 f_2$ means that factors F_1 and F_2 are at the second level and the remaining $m - 2$ factors are at the first level. The treatment which contains all factors at the first level is denoted by the symbol 1. When they refer to numbers, the letters F_i, $F_i F_j$, $F_i F_j F_k$, etc. will represent, respectively, the main effect of F_i, the first order interaction of F_i and F_j, the second order interaction of F_i, F_j and F_k, etc.

It is very well known that each interaction can be expressed as a linear contrast of all treatments. For example, a mathematical expression for representing the contrasts in a general 2^m factorial experiment is $(f_1 \pm 1)(f_2 \pm 1)\dots(f_m \pm 1)$, where "+" is taken for absence and "-" is taken for presence of the corresponding letter in the interaction under consideration,

Let f denote the column vector of all treatments, where $f' = [1; f_1, f_2 \dots; f_1 f_2 \dots; \dots; \dots f_1 f_2 \dots f_m]$. Let F denote the column vector of F's in the same order where the first position represents the mean, μ. Then we may represent the above mentioned contrasts in matrix notation as:

$$F = \delta' f \tag{1.4.1}$$

where $\mathcal{S}'$ is a $2^m \times 2^m$ matrix of plus and minus ones, and any two rows of $\mathcal{S}'$ are orthogonal. Since $\frac{1}{2^m}\mathcal{S}\mathcal{S}' = I_{2^m}$, multiplying both sides of (1.4.1) by $\frac{1}{2^m}\mathcal{S}$ gives

(1.4.2) $$f = \frac{1}{2^m}\mathcal{S}F \, .$$

The matrix $\mathcal{S}$ is sometimes referred to as the Effect Matrix

Let $A = (m,N,2,2)$ be a PBA with index set $\{\mu_0,\mu_1,\mu_2\}$. Each column of A represents a treatment. For a given column, if there is a one in row i, then the corresponding treatment will have factor i at the second level, and if there is a zero in row i, factor i will be at its first level. Let y be an N rowed column vector, where the i^{th} entry in y represents the yield of the treatment which corresponds to the i^{th} column of A.

Using A, we wish to estimate the mean and the m main effects under the assumption that no interactions of two or more factors are present. Thus $F' = (\beta', I_0')$ where $\beta' = [\mu; F_1,F_2,\ldots,F_m]$ and I_0 is a vector of all zeros. Let $\mathcal{S}_0$ be the matrix which contains the first $m+1$ columns of $\mathcal{S}$, then from equation (1.4.2) we have

$$f = \frac{1}{2^m}\mathcal{S}_0\beta \, .$$

It is seen from this equation that each entry in f corresponds to a row of $\mathcal{S}_0$. That is, 1 corresponds to the first row of $\mathcal{S}_0$, f_1 to the second row, and so forth. Using this correspondence, we generate an $(N \times m+1)$ matrix X.

Consider the j^{th} column of A. Then this column corresponds to a particular treatment, t (say). We take as the j^{th} row of X the row in $\mathcal{S}_0$ which corresponds to treatment t.

EXAMPLE: Let

$$A = \begin{array}{cccccc} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ (f_3) & (f_2) & (f_2f_3) & (f_1) & (f_1f_3) & (f_1f_2) \end{array}$$

where the letters in parentheses are the treatments represented by the columns. Then $m = 3$, $f' = [1;\ f_1,f_2,f_3;\ f_1f_2,f_1f_3,f_2f_3;\ f_1f_2f_3]$, $\beta' = [\mu;\ F_1,F_2,F_3]$,

$$\mathcal{S}_0 = \begin{array}{c} \begin{bmatrix} 1 & -1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix} \end{array} \begin{array}{l} (1) \\ (f_1) \\ (f_2) \\ (f_3) \\ (f_1f_2) \\ (f_1f_3) \\ (f_2f_3) \\ (f_1f_2f_3) \end{array} \quad , \text{ and}$$

$$X = \begin{bmatrix} 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 \\ 1 & 1 & 1 & -1 \end{bmatrix} \begin{array}{l} (f_3) \\ (f_2) \\ (f_2f_3) \\ (f_1) \\ (f_1f_3) \\ (f_1f_2) \end{array} \quad , \text{ where the}$$

treatments are given next to their corresponding row in $\mathcal{S}_0$ and X.

Let us make the assumption that the application of the treatments represented by the columns of A is done using a completely randomized design with one replication per column. Thus we may assume that there are no block effects. We further assume that $Var(y) = \sigma^2 I_N$ and that the effects are additive. The construction of X then gives that the expected value of y, written E(y), is equal to $\frac{1}{2^m} X\beta$.

The normal equations are

$$\frac{1}{2^m} X'X\hat{\beta} = X'y, \tag{1.4.3}$$

so that if $(X'X)^{-1}$ exists, the least-square estimates are given by

$$\hat{\beta} = 2^m(X'X)^{-1}X'y. \tag{1.4.4}$$

Consider $X'X$. We may label the rows and columns of $X'X$ with the elements of β taken in order. Thus the first row and first column will be labeled with μ, the second row and the second column with F_1, and so forth. Let $x(e_1,e_2)$ denote the element in $X'X$ which stands at the intersection of the row corresponding to e_1 and the column corresponding to e_2, where e_1 and e_2 are two not necessarily distinct elements of β. Following the proof of Theorem 3.1 in Bose and Srivastava (1964) one can show that

$$x(\mu,\mu) = x(F_i,F_i) = N$$

$$x(\mu,F_i) = \mu_2 - \mu_0$$

$$x(F_i,F_j) = \mu_0 - 2\mu_1 + \mu_2 \qquad i \neq j.$$

Let $a = \mu_2 - \mu_0$ and $b = \mu_0 - 2\mu_1 + \mu_2$, then

$$X'X = \begin{bmatrix} N & a & \cdots & & a \\ a & N & b & \cdots & b \\ \cdot & b & \cdot & & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & & \cdot & b \\ a & b & \cdots & b & N \end{bmatrix} \quad (m+1 \times m+1) \quad .$$

It is a straightforward calculation to find $(X'X)^{-1}$. The interested reader may check to see that $(X'X)^{-1}$ is given by the following.

$$(X'X)^{-1} = \begin{bmatrix} \frac{1}{N}\left(1 + \frac{a^2}{N}(mq + m(m-1)r)\right) & -\frac{a}{N}(q + (m-1)r)\, j_m' \\ & \\ -\frac{a}{N}(q + (m-1)r)\, j_m & (q-r)I_m + rJ_m \end{bmatrix}$$

where j_m is an m rowed column vector of ones

J_m is an $m \times m$ matrix of ones

I_m is the $m \times m$ identity matrix

$$q = \frac{N^2 - a^2(m-1) + (m-2)Nb}{(N^2 - a^2m + (m-1)Nb)(N-b)} \qquad \text{and}$$

$$r = -\frac{(Nb - a^2)}{(N^2 - a^2m + (m-1)Nb)(N-b)} .$$

The above can also be used to obtain S_e^2, the sum of squares due to error. Indeed

$$S_e^2 = y'y - y'X\hat{\beta}. \qquad (1.4.5)$$

The number of degrees of freedom for error is $N - (m+1)$. The expressions (1.4.3) and (1.4.5) can be used, for example, to carry out t-tests for hypotheses that any individual effect is zero.

CHAPTER II

SOME BOUNDS ON THE MAXIMUM NUMBER OF CONSTRAINTS

2.1 DEFINITION AND NOTATION

DEFINITION 2.1.1: Let $A = (a_{ij})$ be an $m \times N$ matrix, where the elements a_{ij} of A are symbols $0,1,2,\ldots,s-1$. Consider the s^t $(1 \times t)$ vectors, $X' = (x_1,\ldots,x_t)$, which can be formed where $x_i = 0,1,\ldots,s-1$; $i = 1,\ldots,t$, and associate with each $(t \times 1)$ vector X a positive integer $\lambda(x_1,\ldots,x_t)$, which is invariant under permutations of $(x_1,\ldots,x_t)$. If for every t-rowed submatrix of A the s^t distinct $(t \times 1)$ vectors X occur as columns $\lambda(x_1,\ldots,x_t)$ times, then the matrix A is called a partially balanced array (PBA) of strength t in N assemblies with m constraints (factors), s symbols (levels) and the specified $\lambda(x_1,\ldots,x_t)$ parameters. When $\lambda(x_1,\ldots,x_t) = \lambda$ for all $(x_1,\ldots,x_t)$, A is called an orthogonal array of index λ.

The set of all $\lambda(x_1,\ldots,x_t)$'s of an array of strength t in s symbols will be called the index set of the array and will be denoted by $\Lambda_{s,t}$. The array A will be represented as the PBA (m,N,s,t) with index set $\Lambda_{s,t}$.

In view of the fact that $\lambda(x_1,\ldots,x_t)$ is invariant under permutations of $(x_1,\ldots,x_t)$, we will denote by $\lambda^{i_1,i_2,\ldots,i_r}_{x_1,x_2,\ldots,x_r}$ the number of repetitions of a fixed column of any $t \times N$ subarray of A, where the column contains i_1 x_1's, i_2 x_2's,... and i_r x_r's.

($x_j = 0,1,\ldots,s-1$, $\Sigma_{j=1}^{r} i_j = t$, $r = \min\{s,t\}$).

In case $s = 2$, we will denote λ_0^t by $\mu_0^{(t)},\ldots,\lambda_{0,1}^{t-i,i}$ by $\mu_i^{(t)},\ldots$, and λ_1^t by $\mu_t^{(t)}$. Where no ambiguity can arise, we will omit the superscript t from $\mu_i^{(t)}$ and write simply μ_i. Clearly, μ_i is the number of times a fixed column containing i ones occurs in any $t \times N$ submatrix of A. Finally, we will refer to a $t \times 1$ column as a t-tuple.

In view of the above we have

DEFINITION 2.1.2: (i) Let $r = \min\{s,t\}$, then

$$\Lambda_{s,t} = \{\lambda_{x_1,\ldots,x_r}^{i_1,\ldots,i_r} \,|\, x_j = 0,1,\ldots,s-1,\ i_j = 0,1,\ldots,t, \text{ where}$$

$$j = 1,\ldots,r \text{ and } \Sigma_{j=1}^{r} i_j = t.\}$$

$$\text{(ii)}\quad \Lambda_{2,t} = \{\mu_i \,|\, i = 0,\ldots,t\}\ .$$

2.2 PROPERTIES

PROPERTY 2.2.1: Let $|\Lambda_{st}|$ denote the number of elements in $\Lambda_{s,t}$. Then, for a PBA of strength t in s symbols,

$$|\Lambda_{s,t}| = \binom{s+t-1}{t}\ .$$

Proof: The number of elements in $\Lambda_{s,t}$ corresponds to the number of distinct combinations of s elements taken t at a time, where an element may be repeated $0,1,\ldots,t$ times in a given combination, and order is not a factor. Let the s elements correspond to s cells and the t possible places in a t-tuple to t indistinguishable objects. Then a distinct "t-combination" can be formed by placing zero, one, or more objects into each of the s cells (i.e. corresponding zero, one, or more of the t places available to each of the s elements). We can represent s cells by the spaces between $s+1$ bars. Such a representation must start and finish with a bar. Thus we have $s-1$ bars and t objects to position. In other words, we have $s-1+t$ spaces to fill with $s-1$ bars and t objects. This can be done in $\binom{s+t-1}{t}$ ways.

<u>COROLLARY</u>: $|\Lambda_{2,x+1}| = |\Lambda_{x,3}|$.

<u>Proof</u>: $|\Lambda_{2,x+1}| = \binom{x+2}{2} = \binom{x+2}{x} = |\Lambda_{x,3}|$.

<u>PROPERTY 2.2.2</u>: Let A be a PBA (m,N,s,t) with index set $\Lambda_{s,t}$. Any subarray of A, (m',N,s,t) with m' rows, where $t \le m' < m$, is a PBA with index set $\Lambda_{s,t}$. Thus, if the PBA (m',N,s,t) does not exist, then the PBA (m,N,s,t) cannot exist.

<u>Proof</u>: This follows directly from the definition.

<u>PROPERTY 2.2.3</u>: Let A be a PBA (m,N,s,t) of strength t with index set $\Lambda_{s,t}$. Then A is a PBA $(m,N,s,t-1)$ of strength $t-1$, where

(i) if $s < t$,

$$\lambda^{(t-1)i_1,\ldots,i_s}_{x_1,\ldots,x_s} = \lambda^{(t)i_1+1,\ldots,i_s}_{x_1,\ldots,x_s} + \lambda^{(t)i_1,i_2+1,\ldots,i_s}_{x_1,x_2,\ldots,x_s} +\ldots+ \lambda^{(t)i_1,\ldots,i_s+1}_{x_1,\ldots,x_s} \quad ; \text{ or}$$

(ii) if $s \ge t$,

$$\lambda^{(t-1)i_1,\ldots,i_{t-1}}_{x_1,\ldots,x_{t-1}} = \lambda^{(t)i_1+1,\ldots,i_{t-1}}_{x_1,\ldots,x_{t-1}} +\ldots+ \lambda^{(t)i_1,\ldots,i_{t-1}+1}_{x_1,\ldots,x_{t-1}} +\ldots+ \lambda^{(t)i_1,\ldots,i_{t-1},1}_{x_1,\ldots,x_{t-1},x_t} +\ldots+ \lambda^{(t)i_1,\ldots,i_{t-1},1}_{x_1,\ldots,x_{t-1},x_s},$$

where for $j = t,t+1,\ldots,s$ $x_j \notin \{x_1,\ldots,x_{t-1}\}$.

Proof: It follows directly from the definition that A is also a PBA of strength $t-1$. Thus, we have only to show (i) and (ii). It is sufficient to show (ii) since (i) would follow from a similar argument.

Consider a $(t-1 \times N)$ subarray of A. A $(t-1)$-tuple containing i_1 x_1's, i_2 x_2's,..., and i_{t-1} x_{t-1}'s in fixed positions can occur as a result of deleting an x_j $(j = 1,\ldots,t-1)$ from a t-tuple containing the i_1 x_1's, i_2 x_2's,..., and i_{t-1} x_{t-1}'s in the fixed positions and x_j in the t^{th} position. Also a $(t-1)$-tuple could occur by deleting an x_j $(j = t,\ldots,s)$ from a t-tuple with i_1 x_1's,..., and i_{t-1} x_{t-1}'s in the fixed positions and x_j in the t^{th} position. Thus (ii) follows.

COROLLARY 1: Let A be a PBA $(m,N,2,t)$ with $\Lambda_{2,t} = \{\mu_i^{(t)} \mid i = 0,\ldots,t\}$. Then A is a PBA of strength $t-1$, where $\mu_i^{(t-1)} = \mu_i^{(t)} + \mu_{i+1}^{(t)}$; $i = 0,\ldots,t-1$.

COROLLARY 2: Let A be a PBA $(m,N,2,t)$. Then A is a PBA of strength 2, where

$$\mu_i^{(2)} = \Sigma_{j=i}^{t+i-2} \binom{t-2}{j-i}\mu_j^{(t)} \qquad i = 0,1,2 .$$

Proof: Apply Corollary 1 repeatedly.

Consider a PBA (m,N,s,t). Let N_i $(i = 0,1,\ldots,s-1)$ stand for the number of i's in a given row of A. Then N_i is independent of the row of A which we choose. This follows from the repeated application of Property (2.2.3).

<u>COROLLARY 3</u>: Let A be a PBA $(m,N,2,t)$ with index set $\Lambda_{2,t}$, then

(i) $N_0 = \Sigma_{j=0}^{t-1} \binom{t-1}{j}\mu_j^{(t)}$

(ii) $N_1 = \Sigma_{j=1}^{t} \binom{t-1}{j-1}\mu_j^{(t)}$

(iii) $N = \Sigma_{j=0}^{t} \binom{t}{j}\mu_j^{(t)}$.

<u>Proof</u>: (i) Consider a PBA in two symbols of strength two. Then the zeros in a given row must occur in 2-tuples with either a zero or a one. There are μ_0 $\binom{0}{0}$'s and μ_1 $\binom{0}{1}$'s so that $N_0 = \mu_0 + \mu_1$. In view of Corollary 2, for a PBA of strength t in two symbols

$$N_0 = \mu_0^{(2)} + \mu_1^{(2)}$$

$$= \Sigma_{j=0}^{t-2} \binom{t-2}{j}\mu_j^{(t)} + \Sigma_{j=1}^{t-1} \binom{t-2}{j-1}\mu_j^{(t)}$$

$$= \mu_0^{(t)} + \Sigma_{j=1}^{t-2} [\binom{t-2}{j} + \binom{t-2}{j-1}]\mu_j^{(t)} + \mu_{t-1}$$

$$= \mu_0^{(t)} + \Sigma_{j=1}^{t-2} \binom{t-1}{j}\mu_j^{(t)} + \mu_{t-1}^{(t)}$$

$$= \Sigma_{j=0}^{t-1} \binom{t-1}{j}\mu_j^{(t)} .$$

(ii) As in (i) we see that

$$N_1 = \mu_1^{(2)} + \mu_2^{(2)}$$

$$= \Sigma_{j=1}^{t-1} \binom{t-2}{j-1}\mu_j^{(t)} + \Sigma_{j=2}^{t} \binom{t-2}{j-2}\mu_j^{(t)}$$

$$= \Sigma_{j=1}^{t} \binom{t-1}{j-1}\mu_j^{(t)} .$$

(iii) Since $N = N_0 + N_1$, we have

$$N = \Sigma_{j=0}^{t-1} \binom{t-1}{j}\mu_j^{(t)} + \Sigma_{j=1}^{t} \binom{t-1}{j-1}\mu_j^{(t)}$$

$$= \mu_0^{(t)} + \Sigma_{j=1}^{t-1} [\binom{t-1}{j} + \binom{t-1}{j-1}]\mu_j^{(t)} + \mu_t^{(t)}$$

$$= \mu_0^{(t)} + \Sigma_{j=1}^{t-1} \binom{t}{j}\mu_j^{(t)} + \mu_t^{(t)}$$

$$= \Sigma_{j=0}^{t} \binom{t}{j}\mu_j^{(t)} .$$

Note that the well known relation, $\binom{a}{b} + \binom{a}{b-1} = \binom{a+1}{b}$, was used throughout the above proof.

The above corollary can be generalized to PBA's with $s > 2$ symbols. However, the notation becomes overwhelming rather quickly. Consider the PBA $(m,N,s,2)$ and let $\lambda_{i,j}^{1\ 1} = \lambda_{i,j}$ $i,j = 0,1,\ldots,s-1$. Then

$$N_i = \Sigma_{j=0}^{s-1} \lambda_{ij} ,$$

since i occurs in a row in common with $j = 0,1,\ldots,s-1$ in some other row λ_{ij} times. Also,

$$N = \Sigma_{i=0}^{s-1} N_i = \Sigma_{i=0}^{s-1} \Sigma_{j=0}^{s-1} \lambda_{ij} .$$

To find N_i and N for an array with s symbols and of strength $t > 2$, one would need to apply Property 2.2.3, which is already quite involved.

2.3 DIOPHANTINE EQUATIONS

In this section we will be concerned with a set diophantine equations, which form a set of necessary conditions for the existence of PBA's.

These equations are given by

LEMMA 2.3.1: In the PBA $(m,N,2,t)$ with index set $\{\mu_0,\ldots,\mu_t\}$, let $n_j^{(i)}$ be the number of i-dimensional columns which contain exactly j ones, $i = t,\ldots,m$, $j = 0,\ldots,i$. Then

$$\binom{i}{t}\binom{t}{\ell}\mu_\ell = \Sigma_{j=\ell}^{i-t+\ell} \binom{j}{\ell}\binom{i-j}{t-\ell}n_j^{(i)} ,$$

where $\ell = 0,1,\ldots,t$.

Chopra (1967) has given a very simple straightforward proof for the above when $i = m$. The proof also applies when $t \le i < m$, so that no further work is needed. We shall be concerned with the above equations in a slightly different form and in less generality, so we give a proof to the following corollary.

COROLLARY 2.3.2: In the PBA $(m,N,2,2)$ with index set $\{\mu_0,\mu_1,\mu_2\}$, the following are true.

(i) $\Sigma_{j=0}^{i} \binom{j}{2}n_j^{(i)} = \binom{i}{2}\mu_2$;

(ii) $\Sigma_{j=0}^{i} jn_j^{(i)} = i(\mu_1 + \mu_2)$;

(iii) $\Sigma_{j=0}^{i} n_j^{(i)} = \mu_0 + 2\mu_1 + \mu_2$, $i = 2,\ldots,m$.

Proof: (i) We would like to count all of the possible 2-tuples, $\binom{1}{1}$, which can occur in i rows of A. We can choose two of the i rows in $\binom{i}{2}$ ways, and since the number of $\binom{1}{1}$'s in these two rows is μ_2, the total number of $\binom{1}{1}$'s is $\binom{i}{2}\mu_2$. On the other hand, a column containing j ones will contribute

$\binom{j}{2}$ $\binom{1}{1}$'s to the total. Noting that the number of columns containing j ones is $n_j^{(i)}$, we see that the total number of $\binom{1}{1}$'s is also equal to $\Sigma_{j=2}^{i} \binom{j}{2} n_j^{(i)} = \Sigma_{j=0}^{i} \binom{j}{2} n_j^{(i)}$. Thus,

$$\Sigma_{j=0}^{i} \binom{j}{2} n_j^{(i)} = \binom{i}{2}\mu_2 .$$

(ii) This follows in the same manner as (i), where both sides are equal to the total number of ones in i rows of A.

(iii) This follows, since both sides are equal to N, the number of columns of A.

As an example of the usefulness of the above equations consider the following.

<u>THEOREM 2.3.3</u>: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0,\mu_1,\mu_2\}$. For $m \geq 3$ we have

(i) $\mu_1 \leq \mu_0 + \mu_2$ with equality if and only if $n_0^{(3)} = n_3^{(3)} = 0$.

(ii) Let $\mu_1 = \mu_0 + \mu_2$. Then $m \leq 4$ with equality if and only if $\mu_0 = \mu_2$.

(iii) Let $\mu_1 = \mu_0 + \mu_2$ and $m = 4$. Then A is a PBA of strength 3 with index set $\{\mu_0^{(3)} = 0, \mu_1^{(3)} = \mu_0^{(2)}, \mu_2^{(3)} = \mu_0^{(2)}, \mu_3^{(3)} = 0\}$.

(iv) $\mu_1 = \mu_2$ if and only if $n_1^{(3)} = 3n_3^{(3)}$. $\mu_0 = \mu_1$ if and only if $n_2^{(3)} = 3n_0^{(3)}$.

<u>Proof</u>: (i) By Corollary 2.3.2, for $i = 3$ we see

$$n_2 + 3n_3 = 3\mu_2$$
$$n_1 + 2n_2 + 3n_3 = 3\mu_1 + 3\mu_2$$
$$n_0 + n_1 + n_2 + n_3 = \mu_0 + 2\mu_1 + \mu_2 .$$

Thus (1)
$$n_2 = 3\mu_2 - 3n_3$$
$$n_1 = 3\mu_1 + 3\mu_2 - 2n_2 - 3n_3$$

$$(2) \qquad = 3\mu_1 - 3\mu_2 + 3n_3$$

$$n_0 = \mu_0 + 2\mu_1 + \mu_2 - n_1 - n_2 - n_3$$

$$(3) \qquad = \mu_0 - \mu_1 + \mu_2 - n_3 \ .$$

Since $n_0 \geq 0$, the last equation gives

$$0 \leq \mu_0 - \mu_1 + \mu_2 - n_3 \ ,$$

so $\mu_1 + n_3 \leq \mu_0 + \mu_2$.

Since $n_3 \geq 0$,

$$\mu_1 \leq \mu_1 + n_3 \leq \mu_0 + \mu_2.$$

Clearly, if $n_0 = n_3 = 0$, $\mu_1 = \mu_0 + \mu_2$. Conversely, suppose $\mu_1 = \mu_0 + \mu_2$. Then, from equation (3) above, $n_0 = -n_3$. Since both n_0 and n_3 are non-negative, it follows that $n_0 = -n_3 = 0$.

(ii) Since $n_0 = n_3 = 0$, any column of A can have at most two zeros and two ones. Thus $m \leq 4$.

Suppose $m = 4$, then by Corollary 2.3.2,

$$12\mu_2 = 2n_2^{(4)} + 6n_3^{(4)} + 12n_4^{(4)} = 2n_2^{(4)} \qquad \text{and}$$

$$4(\mu_1 + \mu_2) = n_1^{(4)} + 2n_2^{(4)} + 3n_3^{(4)} + 4n_4^{(4)} = 2n_2^{(4)} \qquad \text{so that}$$

$$4\mu_1 + 4\mu_2 = 12\mu_2 \qquad \text{or}$$

$$\mu_1 = 2\mu_2 \ .$$

Further, since $\mu_1 = \mu_0 + \mu_2$, it is clear that $\mu_0 = \mu_2$.

Suppose $\mu_0 = c = \mu_2$ so that $\mu_1 = 2c$, then A can be written as the juxtaposition of c arrays of the form

$$B = \begin{matrix} 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \end{matrix}$$

Since B is a PBA, the juxtaposition will be a PBA.

(iii) Since B is a PBA of strength 3 with $\mu_0^{(3)} = 0 = \mu_3^{(3)}$ and $\mu_1^{(3)} = \mu_2^{(3)} = 1$, the juxtaposition of μ_0 B's will be a PBA of strength 3 with the indicated index set.

(iv) That $\mu_1 = \mu_2$ if and only if $n_1^{(3)} = 3n_3^{(3)}$ follows from equation (2) above.

Suppose $n_2^{(3)} = 3n_0^{(3)}$. Then equations (1) and (3) give $3\mu_2 - 3n_3 = 3\mu_0 - 3\mu_1 + 3\mu_2 - 3n_3$ i.e. $\mu_0 = \mu_1$. Now suppose $\mu_0 = \mu_1$, then, by (3), $n_0 = \mu_2 - n_3$. Multiplying both sides by 3 gives $3n_0 = 3\mu_2 - 3n_3 = n_2$, by equation (1).

The equations of Corollary 2.3.2 can be used to find other relations between the elements of the index set. For example, for $m \geq 4$, one can show that $2\mu_1 \leq \min\{3\mu_0 + \mu_2, \mu_0 + 3\mu_2\}$. Furthermore, the equations can be generalized to arrays of strength higher than 2 and similar relations may be found. For example, for $t = 3$ and $m \geq 4$

$$\mu_2^{(3)} \leq \mu_1^{(3)} + \mu_3^{(3)} \quad \text{with equality iff} \quad n_0^{(4)} = n_4^{(4)} = 0$$

$$\mu_1^{(3)} \leq \mu_0^{(3)} + \mu_2^{(3)} \quad \text{with equality iff} \quad n_0^{(4)} = n_3^{(4)} = 0 .$$

And, for $t = 4$ with $m \geq 5$,

$$\mu_1^{(4)} + \mu_3^{(4)} \leq \mu_0^{(4)} + \mu_2^{(4)} + \mu_4^{(4)} \quad \text{with equality iff} \quad n_0^{(5)} = n_5^{(5)} = 0$$

$$\mu_3^{(4)} \leq \mu_2^{(4)} + \mu_4^{(4)} \quad \text{with equality iff} \quad n_2^{(5)} = n_5^{(5)} = 0$$

$$\mu_2^{(4)} \le \mu_1^{(4)} + \mu_3^{(4)} \qquad \text{with equality iff } n_1^{(5)} = n_4^{(5)} = 0.$$

A second example of the usefulness of the diophantine equations of Corollary 2.3.2 is that their solutions are useful in the construction of PBA's.

<u>EXAMPLE 2.3.4</u>: We wish to construct a PBA $(m,5,2,2)$ with index set $\{\mu_0 = 2, \mu_1 = 1, \mu_2 = 1\}$ with as many rows as possible.

(i) Without loss of generality we can write down the first two rows as

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

(ii) For $i = 3$, the equations in Corollary 2.3.2 are

$$3 = n_2 + 3n_3$$

$$6 = n_1 + 2n_2 + 3n_3$$

$$5 = n_0 + n_1 + n_2 + n_3.$$

The solutions are

	n_0	n_1	n_2	n_3
(a)	1	3	0	1
(b)	2	0	3	0 .

The resulting arrays can be written as

(a) $$\begin{matrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \end{matrix}$$ (b) $$\begin{matrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \end{matrix}$$

so that, for example, in (a) there is one column with three ones, three columns with one one, and one column with no ones.

(iii) For $i = 4$, the equations are

$$6 = n_2 + 3n_3 + 6n_4$$

$$8 = n_1 + 2n_2 + 3n_3 + 4n_4$$

$$5 = n_0 + n_1 + n_2 + n_3 + n_4 .$$

The solutions are

	n_0	n_1	n_2	n_3	n_4
(c)	0	4	0	0	1
(d)	1	2	0	2	0 .

Solution (d) is not compatible with solutions (a) and (b), and solution (c) is not compatible with solution (b). But using solution (a) we find

$$
(c) \quad \begin{matrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{matrix}
$$

(iv) For $i = 5$, the equations are

$$
\begin{aligned}
10 &= n_2 + 3n_3 + 6n_4 + 10n_5 \\
10 &= n_1 + 2n_2 + 3n_3 + 4n_4 + 5n_5 \\
5 &= n_0 + n_1 + n_2 + n_3 + n_4 + n_5 \ .
\end{aligned}
$$

The only solution of these equations is $n_0 = 1$, $n_1 = 1$, $n_2 = 1$, $n_3 = 1$, $n_4 = 1$, and $n_5 = 0$. But this is incompatible with (c) above. Thus the maximum possible value of m is 4, and the PBA (4,5,2,2) with index $\{2,1,1\}$ is given by (c).

2.4 BOUNDS - PBA's (m,N,2,2)

Suppose we are given a PBA, A, with index set $\Lambda_{s,t}$. Then the elements of $\Lambda_{s,t}$ are fixed numbers as are N,s, and t. In fact the only parameter which is not completely fixed is m, the number of constraints (rows) of the PBA. This is clear, since if A has m rows, the PBA obtained from A by deleting the last row is a PBA with the same N,s,t, and $\Lambda_{s,t}$ (see Property 2.2.2). Moreover, it may be possible to add an $m+1^{st}$ row to A and obtain a PBA with the same N,s,t, and $\Lambda_{s,t}$. The problem is to determine the conditions under which this can be done. A partial answer is given in this section by determining some upper bounds on the value of m for certain PBA's.

We start by showing that $m \le N$ for all PBA's.

LEMMA 2.4.1: Let A be a PBA $(m,N,2,t)$ with index set $\Lambda_{2,t} = \{\mu_i^{(t)} | i = 0,\ldots,t\}$. Then the eigenvalues of AA' are $b + mc$ with multiplicity one and b with multiplicity $m - 1$, where $b = \Sigma_{j=1}^{t-1} \binom{t-2}{j-1}\mu_j^{(t)}$ $\quad c = \Sigma_{j=2}^{t} \binom{t-2}{j-2}\mu_j^{(t)}$.

Proof: Recalling Corollary 2 of Property 2.2.3, we recognize b as the number of times the 2-tuple $\binom{0}{1}$ (or $\binom{1}{0}$) occurs in any two rows of A. Likewise, c is the number of times the 2-tuple $\binom{1}{1}$ occurs in any two rows of A.

By simple matrix multiplication we see that

$$(1) \qquad AA' = b\, I_m + c\, J_m ,$$

where I_m is the identity matrix and J_m is an $m \times m$ matrix of all ones. Consider an orthogonal transformation which diagonalizes J_m. Multiplying the right side of (1) by this transformation gives

$$b\, I_m + c \begin{bmatrix} m & 0 & \ldots & 0 \\ 0 & 0 & & . \\ . & & & . \\ . & & & . \\ . & & & . \\ 0 & . & \ldots & 0 \end{bmatrix} \quad m \times m$$

The eigenvalues of this are $b + cm$ with multiplicity one and b with multiplicity $m-1$. Hence, AA' has the indicated eigenvalues with the indicated multiplicities.

It is well known that AA' and $A'A$ have the same non-zero eigenvalues with the same multiplicities. Thus $A'A$, which is $N \times N$, has the above given m positive eigenvalues, and it follows that $N \geq m$.

Now, suppose A is a PBA (m,N,s,t) with index $\Lambda_{s,t}$. We replace each non-zero element in A with a one, so that A becomes a PBA in two symbols, 0 and 1. Then the above gives that $m \leq N$. We have thus shown that $m \leq N$ for any PBA.

<u>THEOREM 2.4.2</u>: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0, \mu_1, \mu_2\}$. If $\mu_1^2 > \mu_0\mu_2$, then

$$m \leq \frac{N\,\mu_1}{\mu_1^2 - \mu_0\mu_2} ,$$

with equality if and only if the number of ones in each column of A is the same.

<u>Proof</u>: Let n_j be the number of columns of A which contain j ones, and let

$$\bar{j} = \frac{1}{N} \Sigma_{j=1}^{m} j n_j \ .$$

Since $n_j \geq 0$ for all j, it follows that

$$0 \leq \Sigma_{j=0}^{m} (j - \bar{j})^2 n_j = \Sigma_{j=1}^{m} j^2 n_j - N(\bar{j})^2 \ .$$

Using Corollary 2.3.2, we see that

$$\bar{j} = \frac{m}{N}(\mu_1 + \mu_2) \qquad \text{and}$$

$$\Sigma_{j=1}^{m} j^2 n_j = m(m-1)\mu_2 + m(\mu_1 + \mu_2) \ .$$

Thus

$$0 \leq m(m-1)\mu_2 + m(\mu_1 + \mu_2) - \frac{1}{N} (m(\mu_1 + \mu_2))^2$$
$$= m^2\mu_2 + m\mu_1 - \frac{m^2}{N} (\mu_1 + \mu_2)^2 \ .$$

Since $m > 0$

$$0 \le m\mu_2 + \mu_1 - \frac{m}{N}(\mu_1 + \mu_2)^2 \qquad \text{and}$$

$$m\left(\frac{1}{N}(\mu_1 + \mu_2)^2 - \frac{N\mu_2}{N}\right) \le \mu_1 \quad ,$$

$$m((\mu_1 + \mu_2)^2 - N\mu_2) \le N\mu_1 \quad ,$$

$$m(\mu_1^2 - \mu_0\mu_2) \le N\mu_1 \; .$$

Since by hypothesis $\mu_1^2 - \mu_0\mu_2 > 0$,

$$m \le \frac{N\mu_1}{\mu_1^2 - \mu_0\mu_2} \; .$$

Suppose that the number of ones in each column of A is the same and equal to k. Then $\bar{j} = \frac{1}{N}\Sigma_{j=1}^{m} jn_j = \frac{1}{N}(k\, n_k) = k$. Furthermore,

$$\Sigma_{j=1}^{m} (j - \bar{j})^2 n_j = (\bar{j} - \bar{j})^2 n_{\bar{j}} = 0 \; ,$$

so that, replacing $\le$ with $=$ in the first part of this proof gives,

$$m = \frac{N\mu_1}{\mu_1^2 - \mu_0\mu_2} \; .$$

Suppose

$$m = \frac{N\mu_1}{\mu_1^2 - \mu_0\mu_2} \; .$$

Then, replacing $\le$ with $=$ in the first part of this proof and following the argument in reverse order, it is clear that

$$\Sigma_{j=0}^{m} (j - \bar{j})^2 n_j = 0.$$

Thus $(j - \bar{j})^2 n_j = 0$ for each $j = 0,\ldots,m$. Since $(j - \bar{j})^2 = 0$ if and only if $j = \bar{j}$, it follows that $n_j = 0$ for all $j \ne \bar{j}$ and that $n_{\bar{j}} = N$, so that the number of ones in each column of A is the same and equal to $\bar{j}$.

EXAMPLE: Consider the PBA $(m,6,2,2)$ with index set $\{\mu_0 = 1, \mu_1 = 2, \mu_2 = 1\}$. Then by the theorem

$$m \le \frac{6(2)}{4-1} = 4.$$

The array is given by

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{matrix} .$$

As mentioned in Section 1.3, a PBA with an equal number of ones per column (say k) is the incidence matrix of a balanced incomplete block (BIB) design.

COROLLARY 2.4.3: A PBA $(m,N,2,t)$ which is also the incidence matrix of a BIB design (m,N,r,k,λ) with $k < m$ has the maximum possible number of rows.

Proof: We need only show that m is a maximum when the PBA is considered to be of strength 2. As an array of strength 2, the PBA has index set $\{\mu_0 = N - 2r + \lambda, \mu_1 = r - \lambda, \mu_2 = \lambda\}$. Thus

$$\begin{aligned} \mu_1^2 - \mu_0\mu_2 &= (r - \lambda)^2 - \lambda(N - 2r + \lambda) \\ &= r^2 - \lambda N \\ &= r^2 - \lambda \frac{rm}{k} \\ &= \frac{r}{k}(rk - \lambda m) \\ &= \frac{N}{m} (r - \lambda) > 0, \end{aligned}$$

where we have used the well known results that for a BIB design, $Nk = rm$ and $r(k-1) = \lambda(m-1)$. Now, since the array has k ones in each column, Theorem 2.4.2 implies that m is the maximum number of rows possible. (Indeed $\dfrac{N\mu_1}{\mu_1^2 - \mu_0\mu_2} = \dfrac{N(r-\lambda)}{\frac{N}{m}(r-\lambda)} = m$).

The corollary implies the following interesting property. Let A be a PBA of strength t in 2 symbols. If, when A is considered to be a PBA of strength 2, $\mu_1^2 \le \mu_0\mu_2$, then A cannot have the same number of ones in each column. This follows, since, if A had k (say) ones in each column, the proof of the corollary would give $\mu_1^2 > \mu_0\mu_2$, a contradiction.

(In the above corollary and property we exclude the case $m = k$, which would mean $\mu_0 = \mu_1 = 0$ and $\mu_2 = N$.)

THEOREM 2.4.4: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0,\mu_1,\mu_2\}$. If $\mu_1^2 = \mu_0\mu_2$, then $m \le N-1$.

Proof: With each column of A we associate a distinct variate. With the column of A, which contains $x_1,\ldots,x_m$ $(x_i = 0,1;\ i = 1,\ldots,m)$ in this order, we associate the variate $f(x_1,\ldots,x_m)$. We consider certain linear functions of these N variates.

Denote by $\sum\limits_N$ the summation over all columns of A. Then, we define the 0^{th} stage function to be

$$\sum_N f(x_1,\ldots,x_m)\ ,$$

the sum of all the variates.

Consider two number, c_0 and c_1, such that

$$(1)\qquad (\mu_0 + \mu_1)c_0 + (\mu_1 + \mu_2)c_1 = 0 \qquad \text{and}$$

$$(2)\qquad \mu_0\, c_0^2 + 2\mu_1\, c_0c_1 + \mu_2\, c_1^2 = 0\ .$$

Choose any row, r, of A. Corresponding to this choice, we can construct the linear function

$$\sum_N c_i(r) f(x_1, \ldots, x_m) \ .$$

In the column of A corresponding to the variate $f(x_1, \ldots, x_m)$, the symbol occurring in row r of the array is x_r (0 or 1). In the linear function constructed, we make the coefficient of $f(x_1, \ldots, x_m)$ equal to c_i if $x_r = i$, $i = 0,1$. The linear functions so defined are called first stage functions. Clearly, there are m first stage functions, one for each row of A.

Provided c_0 and c_1 are not both zero, equation (1) above implies that the first stage functions are orthogonal to the 0^{th} stage functions. Equation (2) implies that each first stage function is orthogonal to each of the other first stage functions. Thus, the m+1 functions defined above are all mutually orthogonal and therefore independent. Since the maximum number of independent linear functions of N variates is N, it follows that $N \geq 1+m$ or $m \leq N-1$.

We now show that not both c_0 and c_1 are zero. Equation (1) gives

$$c_0 = -\frac{\mu_1 + \mu_2}{\mu_0 + \mu_1} c_1 = -K\, c_1 \qquad \text{(say)}$$

Equation (2) gives

$$c_1^2 (K^2 \mu_0) - 2\mu_1 K\, c_1^2 + c_1^2\, \mu_2 = 0$$

$$c_1^2 (K^2 \mu_0 + \mu_2 - 2K\, \mu_1) \qquad = 0 \ .$$

Thus, $c_1 = 0$ as well as $c_0 = 0$, unless $K^2\mu_0 + \mu_2 = 2K\,\mu_1$. But as the following shows $K^2\mu_0 + \mu_2 = 2K\,\mu_1$ if and only if $\mu_1^2 = \mu_0\mu_2$, so that by hypothesis there exist c_0 and c_1 not equal to zero.

$$2\mu_1 K = K^2\mu_0 + \mu_2$$

iff $$2\mu_1 \frac{\mu_1 + \mu_2}{\mu_0 + \mu_1} = \mu_0 \frac{(\mu_1 + \mu_2)^2}{(\mu_0 + \mu_1)^2} + \mu_2$$

iff $$2\mu_1(\mu_1 + \mu_2)(\mu_0 + \mu_1) = \mu_0(\mu_1 + \mu_2)^2 + \mu_2(\mu_0 + \mu_1)^2$$

iff $$2\mu_1^3 + \mu_1^2\mu_0 + \mu_1^2\mu_2 = \mu_0^2\mu_2 + 2\mu_0\mu_1\mu_2 + \mu_0\mu_2^2$$

iff $$\mu_1^2(N) = \mu_0\mu_2(N)$$

iff $$\mu_1^2 = \mu_0\mu_2 \ .$$

<u>THEOREM 2.4.5</u>: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0, 1, \mu_2\}$. Then the maximun value of m is $m' = \max\{\mu_0, \mu_2\} + 2$ and $(m',N,2,2)$ exists.

<u>Proof</u>: Without loss of generality, we write the first two rows of A as

$$\begin{array}{ccccccc} \overbrace{0 \quad \quad 0}^{\mu_0} & 0 & 1 & \overbrace{1 \quad \quad 1}^{\mu_2} \\ 0 \ \ldots \ 0 & 1 & 0 & 1 \ \ldots \ 1 \end{array} .$$

Since $\mu_1 = 1$, the third row must have exactly one one in a column of A which has a zero in the first row of A. Call this column c_1. Likewise, the third row must have exactly one one in a column of A which has a zero in the second row of A. Call this column c_2.

<u>CASE 1</u>: Suppose that c_1 and c_2 are different columns. Then, without loss of generality, we write the first three rows of

A as

$$\begin{array}{ccccccccc} \multicolumn{3}{c}{\overbrace{\hphantom{0\ \cdots\ 0}}^{\mu_0}} & c_1 & c_2 & c_3 & \multicolumn{3}{c}{\overbrace{\hphantom{1\ \cdots\ 1}}^{\mu_2-1}} \\ 0 & & 0 & 0 & 1 & 1 & 1 & & 1 \\ 0 & & 0 & 1 & 0 & 1 & 1 & & 1 \\ 0 & \cdots & 0 & 1 & 1 & 0 & 1 & \cdots & 1 \end{array} \quad .$$

In adding a fourth row, we note that there cannot be more than one one in the first μ_0 columns, else $\mu_1 = 1$ would be contradicted. Suppose there is one one in the first μ_0 columns. Then, since $\mu_1 = 1$, there cannot be a one in column c_1, c_2, or c_3. But this leaves only $\mu_2 - 1$ places in which to place μ_2 ones. Thus, in the fourth row we must put zeros in the first μ_0 columns and one zero and $\mu_2 + 1$ ones in the last $\mu_2 + 2$ columns.

Suppose we place the zero in one of c_1, c_2, or c_3 (c_1 say). Then the number of $\binom{0}{0}$'s occurring in the first and fourth rows of A is $\mu_0 + 1$, a contradiction. It therefore follows that the zero must be in a column which has ones in the first three rows.

Using similar arguments, we can continue to add rows as long as there are columns containing all ones. A will thus have the form of μ_0 columns of zeros and $\mu_2 + 2$ columns containing one 0 and the rest ones, where each row has exactly one zero in the last $\mu_2 + 2$ columns. Clearly the number of rows of A is $\mu_2 + 2$.

<u>CASE 2</u>: Suppose c_1 and c_2 are the same column. Then without loss of generality, we write the first three rows of A as

$$\begin{array}{ccccccccc} \multicolumn{3}{c}{\overbrace{\hphantom{0\ \cdots\ 0}}^{\mu_0-1}} & & & & \multicolumn{3}{c}{\overbrace{\hphantom{1\ \cdots\ 1}}^{\mu_2}} \\ 0 & & 0 & 0 & 0 & 1 & 1 & & 1 \\ 0 & & 0 & 0 & 1 & 0 & 1 & & 1 \\ 0 & \cdots & 0 & 1 & 0 & 0 & 1 & \cdots & 1 \end{array} \quad .$$

Using similar arguments to those used in case 1, we see that A has the first $\mu_0 + 2$ columns with one one and the rest zeros, where each row has exactly one one. The remaining μ_2 columns contain all ones. Clearly, the number of rows of A is $\mu_0 + 2$.

Since cases 1 and 2 are the only ones possible, the theorem follows.

EXAMPLE: Recall Example 2.3.4 of Section 2.3. A is a PBA (4,5,2,2) with index set $\{2,1,1\}$ and was shown to have the form

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 \end{matrix} \quad .$$

THEOREM 2.4.6: Let A be a PBA (m,N,2,2) with index set $\{\mu_0,\mu_1,\mu_2\}$. If ℓ is the number of ones in some column of A, we have

$$0 \le m^2(\mu_0\mu_2 - \mu_1^2 - \mu_2) + m\mu_1(N-1) + 2m\ell(2\mu_1^2 - 2\mu_0\mu_2 + \mu_2 - \mu_1)$$
$$+ \ell^2(4\mu_0\mu_2 - 4\mu_1^2 - \mu_0 + 2\mu_1 - \mu_2) \ .$$

Proof: Without loss of generality, we may assume that the first column of A contains ℓ ones. Consider any two rows of the array. If the first column contains $\binom{0}{0}$, it appears $\mu_0 - 1$ more times. Likewise, a $\binom{0}{1}$ or a $\binom{1}{0}$ appears $\mu_1 - 1$ more times and a $\binom{1}{1}$ appears $\mu_2 - 1$ more times. Since there are ℓ ones in the first column, the number of ways to choose two rows so that the first column contains $\binom{0}{0}$ is $\binom{m-\ell}{2}$. The number of ways to choose two rows so that the first column contains $\binom{0}{1}$ or $\binom{1}{0}$ is $\ell(m-\ell)$, and the number of ways to choose two rows so that the first column contains $\binom{1}{1}$ is $\binom{\ell}{2}$.

Let T be the total number of 2-tuples appearing in columns other than the first which are identical with the corresponding 2-tuple in the first column. Then

$$T = (\mu_0 - 1)\binom{m-\ell}{2} + (\mu_1 - 1)\ell(m - \ell) + (\mu_2 - 1)\binom{\ell}{2} .$$

Let $f(i)$ be the number of columns other than the first which have i coincidences[1] with the first column. Then any column with $i \geq 2$ coincidences with the first column will contribute $\binom{i}{2}$ to T. Thus

$$T = \Sigma_{i=2}^{m} \binom{i}{2} f(i) = \Sigma_{i=0}^{m} \binom{i}{2} f(i) ,$$

so that

$$\Sigma_{i=0}^{m} \binom{i}{2} f(i) = (\mu_0 - 1)\binom{m-\ell}{2} + (\mu_1 - 1)\ell(m - \ell) + (\mu_2 - 1)\binom{\ell}{2} ,$$

or $$\Sigma_{i=0}^{m} i(i - 1)f(i) = m^2(\mu_0 - 1) - m(\mu_0 - 1) + 2m\ell(\mu_1 - \mu_0)$$
$$+ \ell^2(\mu_0 - 2\mu_1 + \mu_2) + \ell(\mu_0 - \mu_2) .$$

In a similar manner we can show

$$\Sigma_{i=0}^{m} i\ f(i) = (\mu_0 + \mu_1 - 1)(m - \ell) + (\mu_1 + \mu_2 - 1)\ell$$
$$= \ell(\mu_2 - \mu_0) + m(\mu_0 + \mu_1 - 1) .$$

Also, clearly

$$\Sigma_{i=0}^{m} f(i) = N - 1 .$$

Let $\bar{f} = \frac{1}{N-1} \Sigma_{i=0}^{m} i\ f(i)$, then, since $f(i) \geq 0$ for all i,

---

[1] For definition see, for example, Bose and Bush (1952).

$$0 \le \Sigma_{i=0}^{m}(i - \bar{f})^2 f(i)$$

$$= \Sigma_{i=0}^{m} i^2 f(i) - (\tfrac{1}{N-1})(\Sigma_{i=0}^{m} i f(i))^2$$

$$= \Sigma_{i=0}^{m} i(i-1)f(i) + \Sigma_{i=0}^{m} i f(i) - \frac{1}{N-1}(\Sigma_{i=0}^{m} i f(i))^2 .$$

Thus, $0 \le (N-1)[m^2(\mu_0 - 1) + m\mu_1 + 2m\ell(\mu_1 - \mu_0)$

$$+ \ell^2(\mu_0 - 2\mu_1 + \mu_2)] - [\ell^2(\mu_2 - \mu_0)^2 +$$

$$2m\ell(\mu_2 - \mu_0)(\mu_0 + \mu_1 - 1) + m^2(\mu_0 + \mu_1 - 1)^2]$$

$$= m^2(\mu_0\mu_2 - \mu_1^2 - \mu_2) + m\mu_1(N - 1)$$

$$+ 2m\ell(2\mu_1^2 - 2\mu_0\mu_2 + \mu_2 - \mu_1)$$

$$+ \ell^2(4\mu_0\mu_2 - 4\mu_1^2 - \mu_0 + 2\mu_1 - \mu_2) .$$

<u>COROLLARY 2.4.7</u>: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0,\mu_1,\mu_2\}$ then:

(i) If $\ell = 0$, $m \le \dfrac{(N-1)\mu_1}{\mu_1^2 - \mu_2(\mu_0 - 1)}$ provided $\mu_1^2 - \mu_2(\mu_0 - 1) > 0$

(ii) If $\ell = m$, $m \le \dfrac{(N-1)\mu_1}{\mu_1^2 - \mu_0(\mu_2 - 1)}$ provided $\mu_1^2 - \mu_0(\mu_2 - 1) > 0$.

In the following, let

$$C_1 = \mu_0\mu_2 - \mu_1^2 - \mu_2$$
$$C_2 = \mu_1(N - 1)$$
$$C_3 = 2(2\mu_1^2 - 2\mu_0\mu_2 + \mu_2 - \mu_1)$$
$$C_4 = 4\mu_0\mu_2 - 4\mu_1^2 - \mu_0 + 2\mu_1 - \mu_2 ,$$

so that $0 \le m^2 C_1 + m\, C_2 + m\ell\, C_3 + \ell^2 C_4$.

EXAMPLE: Let A be the PBA $(m,9,2,2)$ with index $\{3,2,2\}$. Then $C_1 = 0$, $C_2 = 16$, $C_3 = 8$, and $C_4 = 7$, so that

$$0 \le 16m - 8m\ell + 7\ell^2 \quad \text{or}$$

$$m \le \frac{7\ell^2}{8(\ell-2)} ,$$

where since $\mu_2 > 0$, we take ℓ so that $3 \le \ell \le m$. The right side is minimized when $\ell = 4$. Thus, since the maximum value of m is a constant,

$$m \le \frac{7(16)}{8(2)} = 7.$$

Indeed we can express A as follows.

$$A = \begin{matrix} 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{matrix}$$

2.5 BOUNDS - GENERAL CASE

In this section we will be concerned with applying the bounds derived in Section 2.4 to PBA's of strength greater than 2 and with more than 2 symbols.

THEOREM 2.5.1: Let A be the PBA (m,N,s,t) with index set $\Lambda_{s,t}$. Then A contains s PBA's, where each array is of strength $t-1$ in s symbols.

Proof: Choose any row, r, of A. Divide the columns of A into s sets, so that each column in a given set has the same number in row r. Exclude row r, and call the $m-1$ rows of the set, which has j in row r, A_j. Let N_j be the number of j's in row r, then A_j has N_j columns. We show A_j is a PBA $(m-1,N_j,s,t-1)$.

Excluding r, choose any $t-1$ rows of A. By definition, every possible $(t-1)$-tuple must occur in these $t-1$ rows with the j's of row r. Moreover, for A_j,

$$\lambda^{(t-1)i_1,\ldots,i_p}_{\quad x_1,\ldots,x_p} = \lambda^{(t),1,i_1,\ldots,i_p}_{\quad j,x_1,\ldots,x_p} ,$$

where $p = \min\{s,t-1\}$, $x_j = 0,\ldots,s-1$, $i_j = 0,1,\ldots,t-1$, and $\Sigma_{j=1}^{p} i_j = t-1$. (If $j = x_k \in \{x_1,\ldots,x_p\}$,

$$\lambda^{(t)\ 1\ i_1,\ldots,i_p}_{\quad j,x_1,\ldots,x_p} = \lambda^{(t)i_1,\ldots,i_k+1,\ldots,i_p}_{\quad x_1,\ldots,x_k\quad,\ldots,x_p} .)$$

Thus A_j, $j = 0,\ldots,s-1$, is a PBA.

<u>COROLLARY 2.5.2</u>: Let A be the PBA $(m,N,2,t)$ with index set $\{\mu_i | i = 0,\ldots,t\}$. Then A contains 2 PBA's,

$A_0 = (m-1,N_0,2,t-1)$ with index set $\{\eta_i | \eta_i = \mu_i,\ i = 0,\ldots,t-1\}$

and

$A_1 = (m-1,N_1,2,t-1)$ with index set $\{\eta_i | \eta_i = \mu_{i+1},\ i = 0,\ldots,t-1\}$.

<u>THEOREM 2.5.3</u>: Let m_j be the maximum number of rows possible for the PBA A_j of Theorem 2.5.1, $j = 0,\ldots,s-1$. Then

$$m \le \min\,(m_j) + 1 .$$

<u>Proof</u>: This is clear, for, if $m > \min\,(m_j) + 1 = m_i + 1$ (say), then by the proof of Theorem 2.5.1, A_i has at least $m_i + 1$ rows in contradiction to the hypothesis that m_i is the maximum

number of rows of A_i.

<u>EXAMPLE</u>: Consider the PBA $(m,10,3,2)$ with index set $\{2,1,1,2\}$. Then by Corollary 2.5.2

$$A_0 = (m-1,5,2,2) \text{ with index set } \{2,1,1\}$$

$$A_1 = (m-1,5,2,2) \text{ with index set } \{1,1,2\}.$$

The maximum possible number of rows for A_0 is 4. The maximum possible number of rows for A_1 is also 4. Thus, the array can have at most 5 rows. We give the array as follows.

$$\begin{array}{cccccccccc}
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 1 & 0
\end{array}$$

The bounds derived in Section 2.4 can also be useful when directly applied to PBA's of strength 2 in more than 2 symbols. For example, let A be the PBA $(m,20,3,2)$ with $\lambda_{00} = 4$, $\lambda_{01} = 3$, $\lambda_{02} = 3$, $\lambda_{11} = 1$, $\lambda_{12} = 1$ and $\lambda_{22} = 1$. Replacing all nonzero elements in A with 1 gives a PBA $(m,20,2,2)$ with $\mu_0 = 4$, $\mu_1 = 6$, $\mu_2 = 4$. By Theorem 2.4.2 $m \le \frac{20(6)}{36-16} = 6$. Thus A can have at most 6 rows. Indeed A can be expressed as follows:

$$A = \begin{array}{cccccccccccccccccccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\
0 & 0 & 0 & 0 & 1 & 1 & 1 & 2 & 2 & 2 & 0 & 0 & 0 & 1 & 2 & 0 & 0 & 0 & 1 & 2 \\
0 & 2 & 1 & 1 & 1 & 2 & 0 & 2 & 0 & 0 & 0 & 0 & 2 & 0 & 1 & 0 & 2 & 1 & 0 & 0 \\
2 & 1 & 2 & 1 & 0 & 0 & 1 & 2 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 2 & 1 \\
2 & 1 & 1 & 0 & 0 & 0 & 2 & 0 & 1 & 2 & 0 & 0 & 2 & 1 & 0 & 1 & 0 & 2 & 0 & 0 \\
1 & 0 & 0 & 2 & 1 & 2 & 0 & 0 & 1 & 2 & 2 & 1 & 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0
\end{array}$$

CHAPTER III

CONSTRUCTION

We have already seen in Section 2.3 how one might go about constructing a PBA directly. In this chapter we will consider ways of constructing PBA's of strength t in s symbols from arrays of strength $t' \leq t$ in s' symbols, where $s' \leq s$.

3.1 CONSTRUCTION OF PBA'S OF STRENGTH $t + 1$ FROM PBA'S OF STRENGTH t.

THEOREM 3.1.1: Let A be a PBA $(t+1,N,2,t)$ with index set $\Lambda_{2,t} = \{\mu_i | i = 0,\ldots,t\}$. Consider the $\binom{t+1}{k}$ possible distinct $(t+1 \times 1)$ vectors which contain k ones. Then each of these vectors appears as a column of A the same number of times, m_k (say). Moreover,

$$m_k = \mu_k - m_{k+1},$$

where $k = 0,\ldots,t$.

Proof: Let $(x_1,\ldots,x_{t+1})'$ be a column of A containing k ones, where each x_i assumes the value 0 or 1. Let

$$x_i^* = 0 \quad \text{if} \quad x_i = 1$$
$$x_i^* = 1 \quad \text{if} \quad x_i = 0 .$$

Then, if $x_j = 0$, $(x_1,\ldots,x_i,\ldots,x_j^*,\ldots,x_{t+1})$ contains $(k+1)$ ones. Let $n(x_1,\ldots,x_{t+1})$ be the number of times the column containing $(x_1,\ldots,x_{t+1})$ appears. Finally, assume $x_i = 1$, $x_j = 0$.

Consider the t-rowed array formed from A by excluding row j. Then, since A is of strength t,

$$n(x_1,\ldots,x_i,\ldots,x_{j-1},x_{j+1},\ldots,x_{t+1}) = \mu_k .$$

Thus,

$$(1) \qquad n(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_{t+1}) + n(x_1,\ldots,x_i,\ldots,x_j^*,\ldots,x_{t+1}) = \mu_k.$$

Consider the t-rowed array formed from A by excluding row i. Then, since A is of strength t,

$$n(x_1,\ldots,x_{i-1},x_{i+1},\ldots,x_j^*,\ldots,x_{t+1}) = \mu_k.$$

Thus,

$$(2) \qquad n(x_1,\ldots,x_i^*,\ldots,x_j^*,\ldots,x_{t+1}) + n(x_1,\ldots,x_i,\ldots,x_j^*,\ldots,x_{t+1}) = \mu_k.$$

Subtracting (2) from (1) gives

$$n(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_{t+1}) = n(x_1,\ldots,x_i^*,\ldots,x_j^*,\ldots,x_{t+1}).$$

Proceeding in this manner, we can show that every column containing k ones must occur the same number of times, m_k (say). Moreover, from (1) it is clear that

$$m_k + m_{k+1} = \mu_k$$

or

$$m_k = \mu_k - m_{k+1} .$$

<u>PROPERTY 3.1.2</u>: $m_k = \mu_k - m_{k+1}$ for all $k = 0,\ldots,t$ if and only if

$$m_k = \Sigma_{i=0}^{j}(-1)^i \mu_{k+i} + (-1)^{j+1} m_{k+j+1},$$

$k = 0,\dots,t,\ j = 0,\dots,t-k.$

<u>Proof</u>: Setting $j = 0$ in

$$m_k = \Sigma_{i=0}^{j}(-1)^i \mu_{k+i} + (-1)^{j+1} m_{k+j+1} \qquad \text{gives}$$

$$m_k = \mu_k - m_{k+1} .$$

Conversely, $m_k = \mu_k - m_{k+1}$ and $m_{k+1} = \mu_{k+1} - m_{k+2}$ imply

$$m_k = \mu_k - \mu_{k+1} + m_{k+2} .$$

Proceeding in this way we find that

$$m_k = \Sigma_{i=0}^{j}(-1)^i \mu_{k+i} + (-1)^{j+1} m_{k+j+1} ,$$

$k = 0,\dots,t$ and $j = 0,\dots,t-k.$

EXAMPLE: Let A be a PBA $(m,N,2,2)$. Then Theorem 3.1.1 gives that, in every three rows of A, the 3-tuples $\begin{matrix}0\\0\\1\end{matrix}$, $\begin{matrix}0\\1\\0\end{matrix}$, and $\begin{matrix}1\\0\\0\end{matrix}$ each appear the same number of times, m_1 (say). (m_1 can vary depending upon the three rows chosen.) Likewise, in every three rows of A, the 3-tuples $\begin{matrix}1\\1\\0\end{matrix}$, $\begin{matrix}1\\0\\1\end{matrix}$, and $\begin{matrix}0\\1\\1\end{matrix}$ each appear the same number of times, m_2 (say). Moreover, given three rows, Property 3.1.2 gives

$$m_0 = \mu_0 - m_1 = \mu_0 - \mu_1 + m_2 = \mu_0 - \mu_1 + \mu_2 - m_3$$

$$m_1 = \mu_1 - m_2 = \mu_1 - \mu_2 + m_3$$

$$m_2 = \mu_2 - m_3 .$$

Let A be given by

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{matrix} \; .$$

Then for rows 1, 2, and 3 $m_0 = 1, m_1 = 1, m_2 = 1, m_3 = 0$,
but for rows 1, 2, and 4 $m_0 = 0, m_1 = 2, m_2 = 0, m_3 = 1$.
Notice further that the above are the only two possible solutions of

$$m_0 = 2 - m_1$$
$$m_1 = 2 - m_2$$
$$m_2 = 1 - m_3 ,$$

such that $m_i \geq 0$, $i = 0,1,2,3$.

We next give a converse to Theorem 3.1.1.

THEOREM 3.1.3: Given a set $\Lambda = \{\mu_i | i = 0,\dots,t\}$ of positive integers, a PBA $(t+1,N,2,t)$ exists with Λ as its index set if there is a solution to the equations

$$m_i = \mu_i - m_{i+1}$$

for $i = 0,\dots,t$, where m_i is a non-negative integer for all i.

Proof: Without loss of generality, we can write down the first t rows of the array. We add the $t+1^{st}$ row as follows. In the first t rows, there are μ_k columns containing k ones in fixed positions. We put a zero in the $t+1^{st}$ row of m_k of these columns and a one in the $t+1^{st}$ row of the remaining m_{k+1} of these columns. Call the resulting array A.

To show A is a PBA, we must show that each choice of $t-1$ out of the first t rows together with the $t+1^{st}$ row satisfies the property that a column containing k ones occurs μ_k times. Without loss of generality, we show this for the first $t-1$ rows of A together with the $t+1^{st}$ row.

Let $(x_1,\ldots,x_{t-1},x_{t+1})$ contain k ones.

<u>CASE 1</u>: Suppose $(x_1,\ldots,x_{t-1})$ contains $k-1$ ones, then

$$
\begin{aligned}
& n(x_1,\ldots,x_{t-1},0,1) + n(x_1,\ldots,x_{t-1},1,1) \\
&= m_k + m_{k+1} && \text{(by construction)} \\
&= \mu_k && \text{(by hypothesis)} ,
\end{aligned}
$$

where the notation is as used in Theorem 3.1.1. Thus

$$n(x_1,\ldots,x_{t-1},x_{t+1}) = n(x_1,\ldots,x_{t-1},1) = \mu_k.$$

<u>CASE 2</u>: Suppose $(x_1,\ldots,x_{t-1})$ contains k ones, then

$$
\begin{aligned}
& n(x_1,\ldots,x_{t-1},0,0) + n(x_1,\ldots,x_{t-1},1,0) \\
&= m_k + m_{k+1} \\
&= \mu_k .
\end{aligned}
$$

Thus

$$n(x_1,\ldots,x_{t-1},x_{t+1}) = n(x_1,\ldots,x_{t-1},0) = \mu_k .$$

Since cases 1 and 2 exhaust the possible situations, the result follows.

It is interesting that Theorem 3.1.1 and Theorem 3.1.3 taken together give necessary and sufficient conditions for the existence of a PBA of strength t and $t+1$ constraints.

<u>THEOREM 3.1.4</u>: Let S be an ordered set of s elements, $e_0, e_1, \ldots, e_{s-1}$. For any positive integer t, consider the s^t different ordered t-tuples of the elements of S. These can be divided into s^{t-1} sets, each set consisting of s t-tuples and closed under cyclic permutations of the elements of S. Denote these sets by S_i, $i = 1,2,\ldots,s^{t-1}$.

Suppose that it is possible to find a scheme T of m rows with elements belonging to S such that, in every t-rowed submatrix, the number of columns belonging to an S_i is constant and greater than zero, with the restriction that if S_j contains a column which is a rowwise permutation of a column in S_i, then the number of columns occurring in a t-rowed submatrix from S_j is the same as the number of columns occurring in a t-rowed submatrix from S_i. Then, one can use this scheme to construct a PBA, $A = (m,N,s,t)$, with index set $\Lambda_{s,t}$.

<u>Proof</u>: We can define the sets S_i, $i = 1,\ldots,s^{t-1}$, as follows. Consider the s^{t-1} distinct (t-1)-tuples formed from elements of S. Let the first t-tuple of each S_i be $(e, e_{i_1}, \ldots, e_{i_{t-1}})'$, where e is a fixed element arbitrarily chosen from S, and $(e_{i_1}, \ldots, e_{i_{t-1}})'$ is one of the distinct (t-1)-tuples formed from elements of S. The additional $s-1$ t-tuples of each of the sets S_i are obtained from the first by cyclic permutation of the elements of S.

A can now be constructed. Append to the columns of the scheme T all the transformations of these columns consisting of cyclic permutations of the elements of S. Choosing any t rows, we see that each of the possible s^t t-tuples occurs. Furthermore, because of the restriction that, if S_j contains a permutation of a column in S_i, then the same number of columns must occur from each, it follows that a permutation of a t-tuple will occur the same number of times as the t-tuple. Finally, the number of occurrences is independent of the t rows chosen.

<u>EXAMPLE</u>: Let $S = \{0,1\}$ and $t = 3$. The 4 distinct 2-tuples, which can be formed from S are $\binom{0}{0}$ $\binom{0}{1}$ $\binom{1}{0}$ $\binom{1}{1}$. Pick 0 as the element e of the proof, then

$$S_1 = \left\{ \begin{matrix} 0 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 1 \\ 1 \\ 1 \end{matrix} \right\}, \quad S_2 = \left\{ \begin{matrix} 0 \\ 0 \\ 1 \end{matrix}, \begin{matrix} 1 \\ 1 \\ 0 \end{matrix} \right\}$$

$$S_3 = \left\{ \begin{matrix} 0 \\ 1 \\ 0 \end{matrix}, \begin{matrix} 1 \\ 0 \\ 1 \end{matrix} \right\}, \text{ and } S_4 = \left\{ \begin{matrix} 0 \\ 1 \\ 1 \end{matrix}, \begin{matrix} 1 \\ 0 \\ 0 \end{matrix} \right\}.$$

Let

$$T = \begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{matrix},$$

and let s_i be the number of times columns from S_i occur in any three rows of T. Then one can check that

$$s_1 = 1 \quad \text{and} \quad s_2 = s_3 = s_4 = 3.$$

Thus we have a PBA of strength 3 given by

$$A = \begin{matrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 \end{matrix}.$$

A close look at T in the above example reveals that it is a PBA of strength 2. We are thus lead to ask whether it is always true that a PBA of strength 2 forms a scheme for the construction of a PBA of strength 3. This is answered by the following.

THEOREM 3.1.5: Let $t = 2u$. Then a PBA $(m,N,2,t)$ with index set $\{\mu_i^{(t)} | i = 0,\ldots,t\}$ forms a scheme for the construction of a PBA $(m,2N,2,t+1)$ with index set $\{\mu_i^{(t+1)} | i = 0,\ldots,t+1\}$, where

$$\mu_k^{(t+1)} = \mu_{t+1-k}^{(t+1)} = \Sigma_{i=0}^{t-2k} (-1)^i \mu_{k+i}^{(t)} \quad , \ k \le u \ .$$

Proof: A set S_i is composed of vectors

$$v = (x_1,\ldots,x_{t+1})' \qquad \text{and}$$

$$v^* = (x_1^*,\ldots,x_{t+1}^*) \ ,$$

where if v has k ones, then v^* has $t+1-k$ ones.

Suppose v contains k ones, then, by Theorem 3.1.1, v will appear in $t+1$ rows of $(m,N,2,t)$ m_k times, and v^* will appear in those $t+1$ rows m_{t+1-k}. Moreover, by Property 3.1.2

$$m_k = \Sigma_{i=0}^{t-2k} (-1)^i \mu_{k+i}^{(t)} + (-1)^{t-2k+1} m_{t+1-k}, \text{ where } \ 2k \le t \ .$$

Since $t - 2k+1 = 2(u - k) + 1$ is odd, we have

$$m_k + m_{t+1-k} = \Sigma_{i=0}^{t-2k} (-1)^i \mu_{k+i}^{(t)} \ ,$$

which is independent of the $t+1$ rows chosen from $(m,N,2,t)$.

Suppose S_j contains a permutation of v. Call it u. Then u^* is a permutation of v^*, and, since m_k and m_{t+1-k} are independent of order, the number of columns of S_j occurring in any $t+1$ rows of $(m,N,2,t)$ is

$$\Sigma_{i=0}^{t-2k} (-1)^i \mu_{k+i}^{(t)} \quad , \ 2k \le t \ ,$$

the same as for S_i. Thus $(m,N,2,t)$ forms a scheme satisfying Theorem 3.1.4.

It is clear from the construction of $(m,2N,2,t+1)$ that a $(t+1)$-tuple containing k ones will appear $m_k + m_{t+1-k}$ times. Likewise, a $(t+1)$-tuple containing $t+1-k$ ones will appear $m_{t+1-k} + m_k$ times. Thus, where $k \le \frac{t}{2} = u$, we have

$$\mu_k^{(t+1)} = \mu_{t+1-k}^{(t+1)} = \Sigma_{i=0}^{t-2k} (-1)^i \mu_{k+i}^{(t)} .$$

<u>COROLLARY 3.1.6</u>: A PBA $(m,N,2,2)$ with index set $\{\mu_0^{(2)}, \mu_1^{(2)}, \mu_2^{(2)}\}$ forms a scheme for the construction of a PBA $(m+1,2N,2,3)$ with index set $\{\mu_0^{(3)}, \mu_1^{(3)}, \mu_2^{(3)}, \mu_3^{(3)}\}$ if either $\mu_1^{(2)} = \mu_2^{(2)}$ or $\mu_0^{(2)} = \mu_1^{(2)}$. Furthermore, if m' is the maximum number of constraints of $(m,N,2,2)$, then $m'+1$ will be the maximum number of constraints of $(m+1,2N,2,3)$.

<u>Proof</u>: Without loss of generality, we assume $\mu_0^{(2)} = \mu_1^{(2)}$. (If $\mu_1^{(2)} = \mu_2^{(2)}$, we can interchange zeros and ones in $(m,N,2,2)$ and obtain an array where $\mu_0^{(2)} = \mu_1^{(2)}$.)

Let T be the PBA $(m,N,2,2)$ and T^* be the array obtained from T by interchanging zeros and ones (i.e. by a cyclic permutation of the elements of $S = \{0,1\}$). Let TT^* represent the juxtaposition of T and T^*, then, by Theorem 3.1.5, TT^* is a PBA $(m,2N,2,2)$ with index $\{\mu_0^{(3)} = \mu_2^{(2)}, \mu_1^{(3)} = \mu_0^{(2)}, \mu_2^{(3)} = \mu_0^{(2)}, \mu_3^{(3)} = \mu_2^{(2)}\}$.

We add the $m+1^{st}$ row to TT^* by placing a zero in each column of T^* and a one in each column of T. Call the resulting array A.

Consider any two of the first m rows of A together with the $m+1^{st}$ row of A. In T

$$n(0,0) = \mu_0^{(2)},\ n(0,1) = n(1,0) = \mu_0^{(2)},\ \text{and}\ \ n(1,1) = \mu_2^{(2)}.$$

In T^*

$$n(0,0) = \mu_2^{(2)},\ n(0,1) = n(1,0) = \mu_0^{(2)},\ \text{and}\ \ n(1,1) = \mu_0^{(2)}.$$

Thus for the three rows

$$n(0,0,1) = \mu_0^{(2)},\ n(0,1,1) = n(1,0,1) = \mu_0^{(2)},\ n(1,1,1) = \mu_2^{(2)}$$
$$n(0,0,0) = \mu_2^{(2)},\ n(0,1,0) = n(1,0,0) = \mu_0^{(2)},\ n(1,1,0) = \mu_0^{(2)}\ .$$

Since this is independent of the two rows chosen from TT^*, A is a PBA $(m+1,2N,2,3)$.

The remainder of the theorem follows from Theorem 2.5.3.

EXAMPLE: Consider A in the example following Lemma 3.1.3. A is a PBA $(4,20,2,3)$ with index set $\{1,3,3,1\}$. By placing ones under the first 10 columns of A and zeros under the remaining 10 columns, we obtain a PBA $(5,20,2,3)$ with index set $\{1,3,3,1\}$.

Corollary 3.1.6 has application in the following kind of situation. Suppose that, after an experiment has been performed, it becomes desirable to include an additional factor, where the original design was a PBA $(m,N,2,2)$. Then instead of performing an entirely new experiment, we consider the original experiment to be half of an array, $(m+1,2N,2,2)$, and add the remaining half in which the new factor will appear constant at the 1 level.

The problem may arise that, while, in the original experiment, there were no interaction effects, the introduction of a new factor makes this assumption questionable. The corollary shows that the additional treatment combinations may be designed so that the augmented experiment is of strength 3, which will allow estimation of main effects in the presence of first order interactions.

Consider now a PBA A of strength 2 in 3 symbols with index set $\Lambda_{3,2} = \{\lambda_{00},\lambda_{11},\lambda_{22},\lambda_{01},\lambda_{02},\lambda_{12}\}$. If we wish to construct a PBA of strength 3, we must consider the sets S_i of Theorem 3.1.4. These can be given as

$$S_1 = \left\{\begin{matrix} 0 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 1 \\ 1 \\ 1 \end{matrix}, \begin{matrix} 2 \\ 2 \\ 2 \end{matrix}\right\} \qquad S_2 = \left\{\begin{matrix} 0 \\ 0 \\ 1 \end{matrix}, \begin{matrix} 1 \\ 1 \\ 2 \end{matrix}, \begin{matrix} 2 \\ 2 \\ 0 \end{matrix}\right\} \qquad S_3 = \left\{\begin{matrix} 0 \\ 0 \\ 2 \end{matrix}, \begin{matrix} 1 \\ 1 \\ 0 \end{matrix}, \begin{matrix} 2 \\ 2 \\ 1 \end{matrix}\right\}$$

$$S_4 = \left\{\begin{matrix} 0 \\ 1 \\ 0 \end{matrix}, \begin{matrix} 1 \\ 2 \\ 1 \end{matrix}, \begin{matrix} 2 \\ 0 \\ 2 \end{matrix}\right\} \qquad S_5 = \left\{\begin{matrix} 0 \\ 2 \\ 0 \end{matrix}, \begin{matrix} 1 \\ 0 \\ 1 \end{matrix}, \begin{matrix} 2 \\ 1 \\ 2 \end{matrix}\right\} \qquad S_6 = \left\{\begin{matrix} 1 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 2 \\ 1 \\ 1 \end{matrix}, \begin{matrix} 0 \\ 2 \\ 2 \end{matrix}\right\}$$

$$S_7 = \left\{\begin{matrix} 2 \\ 0 \\ 0 \end{matrix}, \begin{matrix} 0 \\ 1 \\ 1 \end{matrix}, \begin{matrix} 1 \\ 2 \\ 2 \end{matrix}\right\} \qquad S_8 = \left\{\begin{matrix} 0 \\ 1 \\ 2 \end{matrix}, \begin{matrix} 1 \\ 2 \\ 0 \end{matrix}, \begin{matrix} 2 \\ 0 \\ 1 \end{matrix}\right\} \qquad S_9 = \left\{\begin{matrix} 0 \\ 2 \\ 1 \end{matrix}, \begin{matrix} 1 \\ 0 \\ 2 \end{matrix}, \begin{matrix} 2 \\ 1 \\ 0 \end{matrix}\right\}$$

Consider three rows of A. Let s_i be the number of times columns of S_i appear in these three rows. Then, since

$$n(0,0,0) + n(0,0,1) + n(0,0,2) + n(1,1,1) + n(1,1,0) + n(1,1,2)$$
$$+ n(2,2,0) + n(2,2,1) + n(2,2,2) = \lambda_{00} + \lambda_{11} + \lambda_{22} ,$$

we see

$$\lambda_{00} + \lambda_{11} + \lambda_{22} = s_1 + s_2 + s_3$$

Similarly $\quad \lambda_{00} + \lambda_{11} + \lambda_{22} = s_1 + s_4 + s_5$

$$= s_1 + s_6 + s_7 ,$$

and

$$\lambda_{01} + \lambda_{02} + \lambda_{12} = s_2 + s_5 + s_8$$
$$= s_2 + s_7 + s_9$$
$$= s_3 + s_4 + s_9$$

$$\begin{aligned}\lambda_{01} + \lambda_{02} + \lambda_{12} &= s_3 + s_6 + s_8 \\ &= s_4 + s_7 + s_8 \\ &= s_5 + s_6 + s_9 \,.\end{aligned}$$

Moreover, these equations are independent of the three rows chosen from A. Solving, we find

(1) $s_2 + s_3 = s_4 + s_5$

(2) $s_2 + s_5 = s_3 + s_6$

(3) $s_2 + s_5 = s_4 + s_7$

(4) $s_4 + s_5 = s_6 + s_7$

(5) $s_2 + s_3 = s_6 + s_7$

(6) $s_2 + s_8 = s_6 + s_9$.

(1) and (2) imply

$$2s_2 = s_4 + s_6 \,.$$

(3) and (4) imply

$$2s_4 = s_2 + s_6 \,.$$

Thus,

(I) $$s_2 = s_4 = s_6 \,.$$

(1), (5) and (I) imply

(II) $$s_3 = s_5 = s_7 \,.$$

And, (6) and (I) imply

(III) $$s_8 = s_9 \,.$$

If A is orthogonal, then

$$s_1 + s_2 + s_3 = s_5 + s_6 + s_9 ,$$

so that

(III') $$s_1 = s_8 = s_9 .$$

We now prove

<u>THEOREM 3.1.7</u>: A PBA $(m,N,3,2)$ with index set $\{\lambda_{00},\lambda_{11},\lambda_{12},\lambda_{01},\lambda_{02},\lambda_{12}\}$ forms a scheme for construction of a PBA $(m,3N,3,3)$ if s_1 and one of s_i, $i = 2,\ldots,7$ are independent of the three rows chosen from A.

<u>Proof</u>: Suppose s_1 and s_2 are constant. Then s_4 and s_6 are also constant, by (I). Furthermore,

$$s_2 + s_5 + s_8 - (s_1 + s_2 + s_3) = \lambda_{01} + \lambda_{02} + \lambda_{12} - (\lambda_0 + \lambda_1 + \lambda_2)$$
$$= K \quad \text{(say)} .$$

Thus,

$$s_8 = s_1 + K ,$$

so s_8 and, by III, s_9 are constant.

Finally,

$$s_3 = \lambda_0 + \lambda_1 + \lambda_2 - (s_1 + s_2) ,$$

so s_3 and, by II, s_5 and s_7 are constant.

Now, by Theorem 3.1.4, the theorem follows.

Note that the resulting array will have

$$\lambda_{000} = \lambda_{111} = \lambda_{222}$$
$$\lambda_{001} = \lambda_{112} = \lambda_{220}$$
$$\lambda_{002} = \lambda_{110} = \lambda_{221}$$

and if A is orthogonal

$$\lambda_{000} = \lambda_{111} = \lambda_{222} = \lambda_{012}.$$

3.2 CONSTRUCTION OF PBA'S OF STRENGTH 2 IN s SYMBOLS FROM PBA'S OF STRENGTH 2 IN TWO SYMBOLS

THEOREM 3.2.1: Consider a PBA which is also the incidence matrix of a BIB design $(v,b,r,k,\lambda = 1)$. The existence of this array is equivalent to the existence of a PBA $A = (m,N,s,2)$ with index set $\Lambda_{s,2}$, where

$$\begin{aligned} m &= r \\ N &= b - r \\ s &= k \\ \lambda_{00} &= b - r - (k - 1)(2r - k - 1) \\ \lambda_{0i} &= r - k \qquad\qquad i = 1,\ldots,k-1 \\ \lambda_{ij} &= 1 \qquad\qquad i,j = 1,\ldots,k-1 \,. \end{aligned}$$

Proof: Let T be the incidence matrix of the BIB design $(v,b,r,k,1)$. Then T is a PBA $(v,b,2,2)$ with index set $\{\mu_0 = b - 2r + 1,\ \mu_1 = r - 1,\ \mu_2 = 1\}$.

Interchanging columns and rows as necessary, we can put T into the following form. The first column contains ones in its first k rows and zeros elsewhere. The second column contains a one in its first row, ones in rows $k+1$ through $2k-1$, and zeros elsewhere. (Since $\mu_2 = 1$, once we put a one in the first row of column two, there can be no ones in rows 2 through k of column two.) In general, for $i = 1,\ldots,r$, the i^{th} column of T contains a one in its first row, ones in rows $i(k-1) - (k-3)$ through $i(k-1) + 1$, and zeros elsewhere.

Let K_i represent the $k-1$ rows from row $i(k-1) - (k-3)$ through row $i(k-1) + 1$, $i = 1,\ldots,r$. Consider column c_j, $j = r+1,\ldots,b$. Since $\mu_2 = 1$, c_j can have at most one one in the rows of K_i, where the other entries are zero. If the one occurs in the first row of K_i, enter a one in row i and column $j-r$ of A. If the one occurs in the second row of K_i, enter a two in row i and column $j-r$, and so on. If no one occurs in the rows of K_i, enter a zero in row i column $j-r$ of A. Clearly, $m = r$, $N = b-r$, and $s = k$, so that we must now show that A is a PBA of strength 2 with the indicated index set.

Consider K_u and K_v, $u,v = 1,\ldots,r$, $u \neq v$. Each row of K_u (K_v) contains $r-1$ ones in the columns from $r+1$ through b. Since $\mu_2 = 1$, each row of K_u must have exactly one of these ones in common with each row of K_v. Thus for rows u and v of A,

$$\lambda_{i,j} = 1 \qquad i,j = 1,\ldots,k-1 \qquad \text{and}$$

$$\begin{aligned}\lambda_{0i} &= r - 1 - (k-1) \\ &= r - k \qquad , \; i = 1,\ldots,k-1 \; .\end{aligned}$$

Moreover,

$$\begin{aligned}\lambda_{00} &= (b - r) - 2\,\Sigma_{i=1}^{k-1} \lambda_{0i} - \Sigma_{i=1}^{k-1} \Sigma_{j=1}^{k-1} \lambda_{ij} \\ &= b - r - 2(k-1)(r-k) - (k-1)(k-1) \\ &= b - r - (k-1)(2r - k - 1) \; .\end{aligned}$$

Since u and v were arbitrary, the λ's are independent of the two rows of A. Thus A is a PBA with the indicated index set.

Given the PBA, A, we can reverse the construction and derive T.

<u>EXAMPLE</u>: Consider the BIB design (13,26,6,3,1), then we write T in the indicated form and A below T, where K_1 contains rows two and three, K_2 contains rows four and five,..., and K_6 contains rows 12 and 13.

$$T = \begin{array}{cccccccccccccccccccccccccc}
1&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
1&0&0&0&0&0&0&0&0&0&0&1&1&1&1&1&0&0&0&0&0&0&0&0&0&0\\
0&1&0&0&0&0&1&0&0&0&0&1&0&0&0&0&1&1&1&0&0&0&0&0&0&0\\
0&1&0&0&0&0&0&1&0&0&0&0&1&0&0&0&0&0&0&1&1&1&0&0&0&0\\
0&0&1&0&0&0&0&1&0&0&0&0&0&1&0&0&1&0&0&0&0&0&1&1&0&0\\
0&0&1&0&0&0&0&0&1&0&0&0&0&0&1&0&0&1&0&1&0&0&0&0&1&0\\
0&0&0&1&0&0&0&0&0&1&0&0&1&0&0&0&0&0&1&0&0&0&1&0&1&0\\
0&0&0&1&0&0&0&0&0&0&1&1&0&0&0&0&0&0&0&1&0&0&0&1&0&1\\
0&0&0&0&1&0&1&0&0&0&0&0&0&0&0&1&0&0&0&0&1&0&0&1&1&0\\
0&0&0&0&1&0&0&0&1&0&0&0&0&1&0&0&0&0&1&0&0&1&0&0&0&1\\
0&0&0&0&0&1&0&0&0&1&0&0&0&0&1&0&1&0&0&0&1&0&0&0&0&1\\
0&0&0&0&0&1&0&0&0&0&1&0&0&0&0&1&0&1&0&0&0&1&1&0&0&0
\end{array}$$

$$A = \begin{array}{cccccccccccccccccccc}
1&1&1&1&1&2&2&2&2&2&0&0&0&0&0&0&0&0&0&0\\
1&2&0&0&0&1&2&0&0&0&1&1&1&2&2&2&0&0&0&0\\
0&1&2&0&0&0&0&1&2&0&1&2&0&2&0&0&1&1&2&0\\
0&0&0&1&2&2&1&0&0&0&0&0&1&2&0&0&1&2&1&2\\
1&0&2&0&0&0&0&2&0&1&0&0&2&0&1&2&0&1&1&2\\
0&0&0&1&2&0&0&0&1&2&1&2&0&0&1&2&2&0&0&1
\end{array}$$

Clearly A is a PBA (6,20,3,2) with $\lambda_{00} = 4$, $\lambda_{0i} = 3$, $i = 1,2$ and $\lambda_{ij} = 1$, $i,j = 1,2$. Moreover, as was shown in the example on page 38, A has the maximum number of rows possible. We state this formally as

<u>THEOREM 3.2.2</u>: The PBA A in Theorem 3.2.1 has the maximum possible number of rows.

Proof: This is clear since T has the maximum possible number of rows as shown in Corollary 2.4.3.

CHAPTER IV

RELATION TO OTHER AREAS OF MATHEMATICS

4.1 TACTICAL CONFIGURATIONS

<u>DEFINITION</u>: A complete $\alpha - t - k - m$ configuration is an arrangement of m elements into blocks of size k so that each set of t elements occurs in exactly α blocks.

Let b be the number of blocks in the configuration. Furthermore, let N_i denote the number of sets each of which contains a fixed set of i elements, where $i = 0,\ldots,t$. Then $N_0 = b$ and $N_t = \alpha$, and in general we have

<u>LEMMA 4.1.1</u>: $$N_i = \frac{\alpha\binom{m-i}{t-i}}{\binom{k-i}{t-i}} \qquad i = 0,\ldots,t\ .$$

<u>Proof</u>: Let S_i be a fixed subset containing i elements. As noted above, if $i = t$, then S_t will appear in exactly α sets. If $i < t$, then S_i can be part of several different sets each having t elements. There are $\binom{m-i}{t-i}$ sets of size t containing S_i. However, there are $\binom{k-i}{t-i}$ sets of size t, containing S_i, in a block of size k $(\geq t)$. Thus S_t occurs in $\dfrac{\alpha\binom{m-i}{t-i}}{\binom{k-i}{t-i}}$ of the blocks of the configuration.

<u>EXAMPLE</u>: Consider the triple system, 1 - 2 - 3 - 7, given by

1	1	1	2	2	3	3
2	4	6	4	5	4	5
3	5	7	6	7	7	6 ,

where the blocks are column vectors, and $\alpha = 1$, $t = 2$, $k = 3$, $m = 7$.
Then

$$N_0 = \frac{\alpha\binom{m}{t}}{\binom{k}{t}} = (1)\,\frac{\binom{7}{2}}{\binom{3}{2}} = 7$$

$$N_1 = 3$$

$$N_2 = 1 .$$

The following theorem is due to Chakravarti (1961). The proof given here is somewhat less complicated than the one given by him.

<u>THEOREM 4.1.1</u>: The existence of an $\alpha - t - k - m$ configuration implies the existence of a PBA $A = (m, N_0, 2, t)$ with index set $\{\mu_i \mid i = 0,\dots,t\}$, where

$$\mu_i = \sum_{j=0}^{t-i} (-1)^j \binom{t-i}{j} N_{i+j} ,$$

provided $\mu_i \geq 0$, $i = 0,1,\dots,t$.

<u>Proof</u>: We form A as follows. Let $a_1, a_2, \dots, a_m$ denote the m elements and $S_1, S_2, \dots, S_{N_0}$ denote the N_0 sets of the configuration. Then, we place a one in the i^{th} row and j^{th} column of A if $a_i \in S_j$. If $a_i \notin S_j$ we place a zero in the i^{th} row and j^{th} column of A, $i = 1,\dots,m$, $j = 1,\dots,N_0$.

Consider any t-rowed submatrix of A, and from its N_0 columns choose a column containing r ones. We show that this column occurs $\mu_r = N_r - \binom{t-r}{1} N_{r+1} + \binom{t-r}{2} N_{r+2} - \dots$ times. By renaming a's as necessary we may assume that the ones in the column correspond to $a_1, \dots, a_r$ and the 0's correspond to $a_{r+1}, \dots, a_t$.

Recall that N_r is the number of sets, S_j, which contain the fixed set $a_1,\ldots,a_r$. Also included in this number are those sets containing $a_1,\ldots,a_r$ and a_i, $i = r+1,\ldots,t$. There are N_{r+1} sets containing $a_1,\ldots,a_r$, a_i, and there are $\binom{t-r}{1}$ a_i's. Thus $N_r - \binom{t-r}{1}N_{r+1}$ is the number of sets containing $a_1,\ldots,a_r$ excluding those which contain $a_1,\ldots,a_r$, a_i $i = r+1,\ldots,t$. However, $\binom{t-r}{1}N_{r+1}$ counts twice those sets containing $a_1,a_2,\ldots,a_r,a_i,a_j$ $i \neq j$ $i,j = r+1,\ldots,t$. It counts three times those sets containing $a_1,\ldots,a_r,a_i,a_j,a_k$, $i \neq j \neq k$ $i,j,k = r+1,\ldots,t$. In general, it counts ℓ times those sets containing ℓ of $a_{r+1},\ldots,a_t$ in addition to $a_1,\ldots,a_r$. Therefore, $N_r - \binom{t-r}{1}N_{r+1}$ excludes one too many of the sets of type $a_1,\ldots,a_r,a_i,a_j$, it excludes two too many of the sets of type $a_1,\ldots,a_r,a_i,a_j,a_k$. In general, it excludes $\ell-1$ too many of the sets containing ℓ of $a_r,\ldots,a_t$ in addition to $a_1,\ldots,a_r$.

There are N_{r+2} sets containing $a_1,\ldots,a_r,a_i,a_j$, $i \neq j$, $i,j = r+1,\ldots,t$, and there are $\binom{t-r}{2}$ possible a_ia_j combinations. However, $\binom{t-r}{2}N_{r+2}$ counts those sets which contain $a_1,\ldots,a_r,a_i,a_j,a_k$ $\binom{3}{2} = 3$ times,... .

Continuing in this way, we find that the number of sets containing $a_1,\ldots,a_r$ and not containing any of $a_{r+1},\ldots,a_t$ is $N_r - \binom{t-r}{1}N_{r+1} + \binom{t-r}{2}N_{r+2} - \ldots + (-1)^{t-r}N_t$, and the result follows.

<u>EXAMPLE</u>: Consider the example following Lemma 4.1.1. The PBA generated by the 1 - 2 - 3 - 7 configuration is

$$\begin{matrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{matrix} ,$$

where

$$\mu_0 = N_0 - 2N_1 + N_2 = 7 - 6 + 1$$
$$= 2$$
$$\mu_1 = N_1 - N_2 = 3 - 1$$
$$= 2$$
$$\mu_2 = N_2 = 1 .$$

<u>COROLLARY 4.1.2</u>: The PBA A in Theorem 4.1.1 has the maximum possible number of rows.

<u>Proof</u>: Since each column of A has k ones, by Theorem 2.4.2 it follows that A has the maximum possible number of rows.

In view of this, it is clear that A is the incidence matrix of a BIB design, (v,b,r,k,λ). In fact,

$$v = m$$
$$b = N_0$$
$$r = N_1$$
$$k = k$$
$$\lambda = N_2 .$$

We now give a converse to Theorem 4.1.1.

<u>THEOREM 4.1.3</u>: Let A be a PBA $(m,N,2,t)$ with index set $\{\mu_i | i = 0,\ldots,t\}$. If A has a constant number of ones per column,

k (say), the existence of A implies the existence of a complete μ_t - t - k - m configuration and a complete μ_0 - t - (m-k) - m configuration.

Proof: Number the rows of A one through m and the columns of A one through N. Then i is in set j, if there is a one in row i and column j, $i = 1,\ldots,m$, $j = 1,\ldots,N$. Since the number of ones in a column is constant and equal to k, we have N sets each containing k elements, where the total number of elements is m.

Consider any fixed set of t elements. This set corresponds to t specific rows of A. Since A is of strength t, the number of columns containing ones in each of these rows is μ_t. Thus, the set of t elements will appear in exactly μ_t sets. It now follows that the above sets form a complete μ_t - t - k - m configuration.

To obtain a complete μ_0 - t - (m-k) - m configuration, interchange the zeros and ones in A and proceed as above.

COROLLARY 4.1.4: Let A be a PBA (m,N,2,2) with index set $\{\mu_0,\mu_1,\mu_2\}$, where $\mu_1 = \mu_2$ (or $\mu_0 = \mu_1$), then, if the number of ones in each column of A is constant, the existence of A implies the existence of a complete μ_0 - 3 - k - 2k configuration, where

$$k = \frac{2\mu_1}{\mu_1 - \mu_0} .$$

Proof: By Corollary 3.1.6, the existence of A implies the existence of a PBA (m+1,2N,2,3), where in this circumstance $\mu_3^{(3)} = \mu_0$.

Let k be the number of ones in a column of A. Then

$$k = \frac{m(\mu_1 + \mu_2)}{N} = \frac{2\mu_1 m}{N} .$$

But

$$m = \frac{N\mu_1}{\mu_1^2 - \mu_0\mu_2} = \frac{N}{\mu_1 - \mu_0} ,$$

so

$$k = \frac{2\mu_1}{\mu_1 - \mu_0} .$$

Let A^* be the array obtained from A by interchanging zeros and ones. Then the number of ones per column in A^*, k^*, is

$$k^* = \frac{m(\mu_0 + \mu_1)}{N} = \frac{\mu_0 + \mu_1}{\mu_1 - \mu_0} .$$

Thus,

$$k^* + 1 = \frac{\mu_0 + \mu_1}{\mu_1 - \mu_0} + \frac{\mu_1 - \mu_0}{\mu_1 - \mu_0} = \frac{2\mu_1}{\mu_1 - \mu_0} = k .$$

By the construction of Corollary 3.1.6, it now follows that (m+1,2N,2,3) has k ones in each column. Thus, by the above theorem, there exists a complete μ_0 - 3 - k - 2k configuration.

<u>COROLLARY 4.1.5</u>: Let A be a PBA (m,N,2,2) with index set $\{\mu_0, \mu_1, \mu_2\}$, where $\mu_0 = \mu_2$. Then, if the number of ones in each column of A is constant, the existence of A implies the existence of a complete $(2\mu_0 - \mu_1)$ - 3 - k - 2k configuration, where

$$k = \frac{\mu_1}{\mu_1 - \mu_2} .$$

<u>Proof</u>: By Theorem 3.1.5, the existence of A implies the existence of (m,2N,2,3) where $\mu_3^{(3)} = 2\mu_0 - \mu_1$. Let k be the number of ones in a column of A. Then

$$k = \frac{m(\mu_1 + \mu_2)}{N}$$

$$= \frac{m(\mu_1 + \mu_2)}{2(\mu_1 + \mu_2)}$$

$$= \frac{m}{2} \quad .$$

Thus, the number of ones in each column of $(m,2N,2,3)$ is $\frac{m}{2}$, and, by Theorem 4.1.3, there exists a complete $(2\mu_0 - \mu_1) - 3 - k - 2k$ configuration. Moreover, since

$$m = \frac{N\mu_1}{\mu_1^2 - \mu_2^2}$$

$$= \frac{2\mu_1(\mu_1 + \mu_2)}{(\mu_1 - \mu_2)(\mu_1 + \mu_2)}$$

$$= \frac{2\mu_1}{\mu_1 - \mu_2} \quad ,$$

$$k = \frac{m}{2}$$

$$= \frac{\mu_1}{\mu_1 - \mu_2} \quad .$$

4.2 GRAPH THEORY

In this section we will be concerned with ordinary graphs which are undirected, without loops or multiple adjacencies, and of a finite order. A graph will be represented by a pair $\{V,M\}$, where V is a set of v vertices and M is the adjacency matrix of the graph.

$$M = (m(x,y))^{v \times v} \quad ,$$

where

$$m(x,y) = \begin{cases} -1 & \text{if } x \text{ and } y \text{ are adjacent} \\ 1 & \text{if } x \text{ and } y \text{ are nonadjacent} \\ 0 & \text{if } x = y \end{cases} ,$$

$x,y \in V$.

Let x be a vertex of a graph. Let $n_1(x)$ be the number of vertices adjacent to x and $n_2(x)$ be the number of vertices nonadjacent to x. Then

$$n_1(x) + n_2(x) = v - 1 ,$$

where v is the number of vertices of the graph. A graph is called REGULAR if $n_1(x)$ (and thus $n_2(x)$) is independent of the choice of x.

Let x and y be any two distinct vertices of a graph. Let

$$h = \begin{cases} 1 & \text{if } x \text{ and } y \text{ are adjacent} \\ 2 & \text{if } x \text{ and } y \text{ are nonadjacent.} \end{cases}$$

Then we define p^h_{ij} as follows. p^h_{11} is the number of vertices which are adjacent to x and to y. p^h_{12} is the number of vertices which are adjacent to x and nonadjacent to y. p^h_{21} is the number of vertices which are nonadjacent to x and adjacent to y. p^h_{22} is the number of vertices nonadjacent to x and to y. A graph is called STRONG if it is nonvoid and noncomplete and if

$$p^1_{12}(x,y) + p^1_{21}(x,y)$$

and

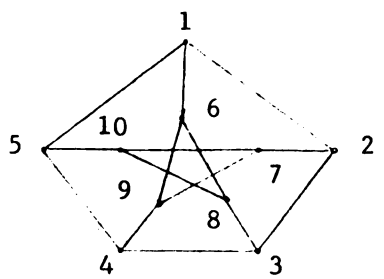
$$p^2_{12}(x,y) + p^2_{21}(x,y)$$

are independent of the choice of x and y. A graph is called STRONGLY REGULAR if it is both strong and regular.

<u>EXAMPLE</u>: PETERSEN GRAPH. Let $V = \{v_i \mid i = 1,\ldots,10\}$ and M be given by

$$M = \begin{matrix} 0 & -1 & 1 & 1 & -1 & -1 & 1 & 1 & 1 & 1 \\ -1 & 0 & -1 & 1 & 1 & 1 & -1 & 1 & 1 & 1 \\ 1 & -1 & 0 & -1 & 1 & 1 & 1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 0 & -1 & 1 & 1 & 1 & -1 & 1 \\ -1 & 1 & 1 & -1 & 0 & 1 & 1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 & 1 & 0 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & 1 & 1 & 1 & 0 & 1 & -1 & -1 \\ 1 & 1 & -1 & 1 & 1 & -1 & 1 & 0 & 1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & -1 & 1 & -1 & -1 & 1 & 0 \end{matrix}$$

Then we may represent the graph pictorially as follows:



This graph is strongly regular, where

$$n_1 = 3 ,$$

$$n_2 = 6 ,$$

$$p^1_{12} + p^1_{21} = 4 , \qquad \text{and}$$

$$p^2_{12} + p^2_{21} = 4 .$$

We next investigate the relation between graphs and PBA's.

<u>DEFINITION</u>: Consider any two columns of a PBA, A. If there are exactly c rows in which each column has a one, we say they have c ones in common. A is called a Quasisymmetric Partially Balanced Array (QPBA) if the number of ones in common to two columns of A can take exactly two values, c_1 and c_2 (say), where $c_1 > c_2$.

<u>EXAMPLE</u>: Let A be given by

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{matrix} \quad .$$

Then the first and second columns have one one in common, but the first and last columns have zero ones in common. In fact, it can be verified that any two columns of A have either zero ones or one one in common. Hence A is a QPBA.

QPBA's are not difficult to find, as the following shows.

<u>THEOREM 4.2.1</u>: Let A be a PBA $(m,N,2,2)$ with index set $\{\mu_0, \mu_1, \mu_2 = 1\}$ and $m < N$. Then A is a QPBA with $c_1 = 1$, $c_2 = 0$.

<u>Proof</u>: Since $\mu_2 = 1$, no two columns of A have more than one one in common. Since $\mu_1 + \mu_2 = \mu_1 + 1 > 1$, there are columns which have one one in common. To show that there are columns which have no ones in common, we suppose there are none.

Let X be any column of A, and suppose that X contains k ones. Corresponding to the first one in X, there are μ_1 columns which have a one, and there are $\mu_0 + \mu_1$ columns which have a zero. Corresponding to the remaining $k-1$ ones in X, there are $(k-1)\mu_1$ columns which have a common one. Since we assume that every column in A has exactly one one in common with X, it follows that

$$(k-1)\mu_1 = \mu_0 + \mu_1 \ .$$

Thus,

$$k = \frac{\mu_0 + 2\mu_1}{\mu_1} = \frac{N-1}{\mu_1}$$

and is independent of the choice of the column X. Therefore A is the incidence matrix of a BIB design $(m, N, r = \mu_1 + 1,\ k = \frac{N-1}{\mu_1}, 1)$, where $m < N$ and

$$\text{(i)} \quad r(k-1) = m-1$$

$$\text{(ii)} \quad Nk = mr$$

$$\text{(iii)} \quad rk-k = N-1 \ .$$

(i) and (iii) give

$$\text{(iv)} \quad r+m = N+k \ .$$

Suppose

$$N = m + C \ ,$$

then (iv) gives

$$r+m = m + C + k$$

so

$$\text{(v)} \quad C = r-k \ .$$

(ii) gives

$$(m+C)k = mr$$

so that

$$\text{(vi)} \quad C = \frac{m(r-k)}{k} \ .$$

Thus, by (v) and (vi),

$$r-k = \frac{m(r-k)}{k}$$

$$m = k ,$$

which implies that each column of A contains all ones. Since $m \geq 2$, this implies $\mu_1 > 1$, a contradiction. Therefore, X must have no ones in common with at least one column of A, from which it follows that A is a QPBA with $c_1 = 1$, $c_2 = 0$.

<u>THEOREM 4.2.2</u>: Let A be a QPBA (m,N,2,2). Then the existence of A implies the existence of a graph on N vertices.

<u>Proof</u>: Since A is a QPBA, any two columns have either c_1 or c_2 ones in common. Choose x and y $(0 < y \leq x)$ so that

$$x+y = c_1 \quad \text{and} \quad x-y = c_2 \ .$$

Also, let k_j be the number of ones in column j of A. Then we write

$$A'A = \begin{bmatrix} k_1 & & \\ & \ddots & \\ & & k_N \end{bmatrix} + x(J_N - I_N) - yM,$$

where J_N is the $N \times N$ matrix of all ones, I_N is the $N \times N$ identity matrix, and M is an $N \times N$ matrix with zeros on the diagonal and plus or minus one off the diagonal. Thus, M is the adjacency matrix of a graph on N vertices, and the result follows.

In view of this theorem and the above definitions of strong graphs and regular graphs, we make the following definitions.

<u>DEFINITIONS</u>: (i) Let X be a column of a QPBA, A. Let $n_1(X)$ be the number of columns of A which have c_1 ones in common with X, and let $n_2(X)$ be the number of columns of A which have c_2 ones in common with X. Then, A will be called REGULAR if $n_1(X)$ and $n_2(X)$ are independent of the choice of X.

(ii) Let X and Y be columns of a QPBA, A. If X and Y have c_1 (c_2) ones in common, let $p^1_{1,2}(X,Y)$ $(p^2_{1,2}(X,Y))$ be the number of columns of A with c_1 ones in common with X and c_2 ones in common with Y. Let $p^1_{2,1}(X,Y)$ $(p^2_{2,1}(X,Y))$ be the number of columns of A with c_2 ones in common with X and c_1 ones in common with Y. Then A will be called STRONG if

$$p^1_{1,2}(X,Y) + p^1_{2,1}(X,Y) \quad \text{and} \quad p^2_{1,2}(X,Y) + p^2_{2,1}(X,Y)$$

are independent of the choice of X and Y.

<u>EXAMPLE</u>: Let

$$A = \begin{matrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{matrix} \ ,$$

then A is strongly regular, where $c_1 = 1$, $c_2 = 0$, $n_1 = 4$, $n_2 = 1$,

$$p^1_{1,2} + p^1_{2,1} = 2 \ ,$$

and

$$p^2_{1,2} + p^2_{2,1} = 0 \ .$$

<u>THEOREM 4.2.3</u>: Let A be a QPBA $(m,N,2,2)$ with index $\{\mu_0,\mu_1,\mu_2\}$. A necessary and sufficient condition for A to be regular is that the number of ones per column be constant. Moreover,

if $\mu_2 \geq 2$, the number of ones per column can take at most two values.

<u>Proof</u>: Let X be a column of A, and suppose X contains k ones. Let n_i be the number of columns which have c_i ones in common with X. Then

$$\text{(i)} \qquad n_1 + n_2 = N-1$$

We can choose a one from X in k ways. Corresponding to this choice, there are $\mu_1 + \mu_2 - 1$ ones in other columns. Thus, there are $k(\mu_1 + \mu_2 - 1)$ pairs of ones, where the first one is from X and the second one is from the same row but a different column. On the other hand, a column with c_i ones in common with X will account for c_i of the above pairs, i = 0,1. Since there are n_i (i = 0,1) such columns, it follows that the number of pairs is $n_1c_1 + n_2c_2$. Thus,

$$\text{(ii)} \qquad n_1c_1 + n_2c_2 = k(\mu_1 + \mu_2 - 1)$$

Substituting (i) into (ii) gives

$$\begin{aligned} k(\mu_1 + \mu_2 - 1) &= n_1c_1 + c_2(N - 1 - n_1) \\ &= n_1(c_1 - c_2) + c_2(N-1) \ . \end{aligned}$$

Thus,

$$\text{(iii)} \qquad n_1 = \frac{k(\mu_1 + \mu_2 - 1) - c_2(N-1)}{c_1 - c_2}$$

and

$$n_2 = N - 1 - n_1 \ .$$

From this it is clear that n_1 and n_2 are independent of X if and only if the number of ones per column of A is constant and equal to k.

We can choose two ones from X in $\binom{k}{2}$ ways. Corresponding to this choice, there are $\mu_2 - 1$ columns which have ones in the same rows as the two chosen ones. Thus there are $\binom{k}{2}(\mu_2 - 1)$ pairs, where the first entry in the pair is two ones from X and the second entry is two ones chosen from the same rows but a different column. On the other hand, a column with c_i ones in common with X will account for $\binom{c_i}{2}$ of the above pairs. Since there are n_i such columns, it follows that the number of pairs is $n_1\binom{c_1}{2} + n_2\binom{c_2}{2}$. Thus,

$$\text{(iv)} \qquad n_1\binom{c_1}{2} + n_2\binom{c_2}{2} = \binom{k}{2}(\mu_2 - 1) .$$

Multiplying both sides by two and substituting using (i) gives

$$n_1c_1(c_1 - 1) + (N - 1 - n_1)c_2(c_2 - 1) = k(k-1)(\mu_2 - 1).$$

Thus, if $c_1 \geq 2$,

$$\text{(v)} \qquad n_1 = \frac{k(k-1)(\mu_2 - 1) - (N-1)c_2(c_2 - 1)}{c_1(c_1 - 1) - c_2(c_2 - 1)} .$$

Subtracting (iii) from (v) and simplifying gives

$$(c_1 - c_2)(k(k-1)(\mu_2 - 1) - (N-1)c_2(c_2 - 1)) =$$
$$(k(\mu_1 + \mu_2 - 1) - c_2(N-1))(c_1(c_1 - 1) - c_2(c_2 - 1)).$$

Or,

$$k^2(c_1-c_2)(\mu_2-1) - k((c_1-c_2)(\mu_2-1) + (\mu_1+\mu_2-1)(c_1(c_1-1) - c_2(c_2-1)))$$
$$- (c_1-c_2)(N-1)c_2(c_2-1) + c_2(N-1)(c_1(c_1-1) - c_2(c_2-1)) = 0 .$$

Or,

$$\text{(vi)} \qquad k^2(\mu_2-1) - k((c_1+c_2)(\mu_1+\mu_2-1) - \mu_1) + c_1c_2(N-1) = 0 .$$

Now, since $\mu_2 \geq 2$, it follows that $c_1 \geq 2$ and that k must satisfy the quadratic equation (vi). Clearly k can take at most two values.

Theorem 4.2.3 says in effect that a QPBA is regular if and only if it is also the incidence matrix of a BIB design with block size k. In this context Seidel (1969) has proved that the array is also strong.

We note that a QPBA which is strong need not be regular. Consider the following array.

$$A = \begin{matrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \end{matrix} .$$

A is not regular, since the number of ones in a column of A is not constant. A is, however, strong, where $c_1 = 1$, $c_2 = 0$, and

$$p^1_{12} + p^1_{21} = 2$$
$$p^2_{12} + p^2_{21} = 2 .$$

4.3 HADAMARD MATRICES:

DEFINITION: A Hadamard Matrix H_p is a square matrix of order p, whose elements are all $+1$ and -1, which satisfies the condition

$$H_p H_p' = p\, I_p ,$$

where I_p is the $p \times p$ identity matrix.

Necessary conditions for the existence of Hadamard Matrices H_p are

$$p = 2 \quad \text{or}$$

$$p \equiv 0 \quad (\text{modulo } 4).$$

Moreover, it has been shown that Hadamard Matrices exist for all $p < 156$ and for infinitely many other values of p subject to the necessary conditions.

<u>EXAMPLE</u>: Using + for +1 and - for -1, we can write a Hadamard Matrix of order 8 as follows.

$$H_8 = \begin{array}{cccccccc}
+ & + & + & + & + & + & + & + \\
+ & + & + & - & + & - & - & - \\
+ & - & + & + & - & + & - & - \\
+ & - & - & + & + & - & + & - \\
+ & - & - & - & + & + & - & + \\
+ & + & - & - & - & + & + & - \\
+ & - & + & - & - & - & + & + \\
+ & + & - & + & - & - & - & +
\end{array}$$

When H_p is written with plus ones in its first column and its first row, we say it is in its normal form. Any Hadamard Matrix can be normalized as follows. If there is a minus one in the first column of any row, multiply that row by minus one. Also, if there is a minus one in the first row of any column, multiply that column by minus one. The result will be a Hadamard Matrix in normal form.

THEOREM 4.3.1: The existence of a Hadamard Matrix of order $p = 4\mu$ implies the existence of the PBA's

(i) $(4\mu-1, 4\mu-1, 2,2)$ with index set $\{\mu-1,\mu,\mu\}$

(ii) $(2\mu\ , 4\mu-2, 2,2)$ with index set $\{\mu-1,\mu,\mu-1\}$

(iii) $(2\mu-1, 4\mu-2, 2,2)$ with index set $\{\mu-2,\mu,\mu\}$.

Moreover, the number of rows in each case is the maximum number possible.

Proof: Without loss of generality, we assume that $H_{4\mu}$ is a Hadamard Matrix in normal form. Let us remove the first row of $H_{4\mu}$ and, in the remaining rows, replace each plus one with a zero and each minus one with a one. Then, it is well known that the resulting array is an orthogonal array $(4\mu-1, 4\mu, 2, 2)$ with index μ. Omitting the first column of this array results in a PBA $(4\mu-1, 4\mu-1, 2, 2)$ with index set $\{\mu-1, \mu, \mu\}$. Call the PBA A_1.

Choose a column, c, from A_1. If a row has a one in c, place that row in an array, A_2 (say). If a row has a zero in c, place that row in an array, A_3 (say). Omitting the ones in column c from A_2 gives a PBA $(2\mu, 4\mu-2, 2, 2)$ with index set $\{\mu-1, \mu, \mu-1\}$. Omitting the zeros in column c from A_3 gives a PBA $(2\mu-1, 4\mu-2, 2, 2)$ with index set $\{\mu-2, \mu, \mu\}$.

That the number of rows is a maximum in each case can be seen using Theorem 2.4.2.

EXAMPLE: Replacing + with 0 and - with 1 in H_8 of the preceeding example and omitting the first row and the first column gives

$$A_1 = \begin{matrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \end{matrix} ,$$

a PBA (7,7,2,2) with index set $\{1,2,2\}$. Letting c be the first column, we find

$$A_2 = \begin{matrix} 1 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 0 & 0 \end{matrix} \quad \text{and}$$

$$A_3 = \begin{matrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \end{matrix} .$$

Omitting the first column from A_2 results in the PBA (4,6,2,2) with index set $\{1,2,1\}$. Omitting the first column from A_3 results in the PBA (3,6,2,2) with index set $\{0,2,2\}$.

It is clear that the PBA's constructed above are also incidences matrices of BIB designs with equal block sizes. Rao (1961) has given a list of BIB designs for $10 \le r \le 15$ together with methods of construction in most cases. Although this list is said to be complete, the design (14, 26, 13, 7, 6) has been omitted. This design is equivalent to the PBA (14,26,2,2) with index $\{6,7,6\}$. Theorem 4.3.1 (ii) with $\mu = 7$ implies the existence of this array and hence the design. We give the design as follows.

2	1	1	1	1	3	4	3	4	1	2	2	5	1	2	3	1	1
3	3	2	2	2	4	5	5	6	5	3	4	6	3	5	8	4	3
4	4	6	7	6	6	7	6	10	7	6	6	7	6	8	9	8	4
5	5	8	8	7	8	9	7	11	10	7	7	9	7	9	11	9	7
6	7	10	9	9	9	10	8	12	11	10	8	12	8	10	12	10	9
9	8	11	11	10	10	11	11	13	13	11	9	13	12	11	13	12	11
11	10	14	12	13	13	12	14	14	14	12	14	14	13	13	14	14	14

1	2	1	2	1	1	2	1
2	3	2	4	3	2	3	4
4	5	3	5	5	3	4	5
5	8	4	7	6	5	7	6
6	10	11	8	9	9	10	8
12	12	12	12	10	13	13	11
14	14	13	13	12	14	14	13

,

where each column represents a block and the numbers represent the application of a particular treatment.

BIBLIOGRAPHY

BIBLIOGRAPHY

Addleman, Sidney (1962), "Augmenting Factorial Plans to Accomodate Additional Two Level Factors", Biometrics pp. 308-322.

Bose, R.C. (1947), "Mathematical Theory of the Symmetrical Factorial Design", Sankhya, Vol. 8 (Part 2), pp. 107-166.

Bose, R.C. and Bush, K.A. (1952), "Orthogonal Arrays of Strength Two and Three", Ann. Math. Statist., Vol. 23, pp. 508-524.

Bose, R.C. and Kishen, K. (1940), "On the Problem of Confounding in the General Symmetrical Factorial Design", Sankhya, Vol. 5, pp. 21-36.

Bose, R.C. and Nair, K.R. (1939), "Partially Balanced Incomplete Block Designs", Sankhya, Vol. 4, pp. 337-372.

Bose, R.C. and Srivastava, J.N. (1964), "Analysis of Irregular Factorial Fractions", Sankhya (series A), Vol. 26, pp. 117-144.

Box, G.E.P. and Hunter, J.S. (1961), "The 2^{k-p} Fractional Factorial Designs". I and II., Technometrics, Vol. 3, pp. 311-351, pp. 449-458.

Chopra, D.V. (1967), "Investigations on the Construction and Existence of Balanced Fractional Factorial Designs of 2^m Series", Ph.D. dissertation, University of Nebraska.

Chakravarti, I.M. (1956), "Fractional Replication in Asymmetrical Factorial Designs and Partially Balanced Arrays", Sankhya, Vol. 17 (part 2), pp. 143-164.

Chakravarti, I.M. (1961), "On Some Methods of Construction of Partially Balanced Arrays", Ann. Math. Statist., Vol. 32, pp. 1181-1185.

Chakravarti, I.M. (1963), "Orthogonal and Partially Balanced Arrays and Their Application in Design of Experiments", Metrika, Vol. 7, pp. 231-243.

Cochran, W.G. and Cox, G.M. (1966), <u>Experimental Design</u>, Wiley and Sons, Inc., Copyright 1957.

Feller, W. (1965), <u>An Introduction to Probability Theory and Its Applications</u>, <u>Volume I</u>, Second Edition, Wiley and Sons, Inc., Copyright 1957.

Fisher, R.A. (1960), *The Design of Experiments*, Seventh Edition, Hafner.

Peng, K.C. (1967), *The Design and Analysis of Scientific Experiments*, Addison-Wesley, Copyright 1967.

Rao, C.R. (1947), "Factorial Arrangements Derivable from Combinatorial Arrangements of Arrays", Supp. of the Jour. of the Roy. Statist. Soc., Vol. 9, pp. 128-139.

Rao, C.R. (1950), "The Theory of Fractional Replication in Factorial Experiments", Sankhya, Vol. 10, pp. 81-86.

Rao, C.R. (1961), "A Study of BIB Designs with Replications 11 to 15", Sankhya, Vol. 23, pp. 117-127.

Rao, C.R. (1968), *Linear Statistical Inference and Its Application*, Wiley and Sons, Inc., Copyright 1965.

Scheffé, H. (1967), *The Analysis of Variance*, Wiley and Sons, Inc., Copyright 1959.

Seidel, J.J. (1969), "Discrete Mathematics", Unpublished class notes given at Michigan State University.

Seiden, E. (1954), "On the Problem of Construction of Orthogonal Arrays", Ann. Math. Statist., Vol. 25, pp. 151-156.

Seiden, E. and Zemach, R. (1966), "On Orthogonal Arrays", Ann. Math. Statist., Vol. 37, pp. 1355-1370.

Sprott, D.A. (1955), "Balanced Incomplete Block Designs and Factorial Configurations", Ann. Math. Statist., Vol. 26, pp. 752-757.

Srivastava, J.N. (1970), "Optimal Balanced 2^m Fractional Factorial Designs", S.N. Roy Memorial Volume, University of North Carolina and Indian Statistical Institute, pp. 689-706.

Srivastava, J.N., "Some General Existence Conditions for Balanced Arrays of Strength t and Two Symbols", To appear in J. Combinatorial Theory.

Srivastava, J.N. and Chopra, D.V. (1971a), "Balanced Optimal 2^m Fractional Factorial Designs of Resolution V, $m \leq 6$", Technometrics, Vol. 13, pp. 257-269.

Srivastava, J.N. and Chopra, D.V. (1971b), "On the Characteristic Roots of the Information Matrix of 2^m Balanced Factorial Designs of Resolution V, With Applications", Ann. Math. Statist., Vol. 42.

Srivastava, J.N. and Chopra, D.V., "Balanced Arrays of Strength 4, 2 Symbols, and $\mu_2 = 1$", To appear in J. Combinatorial Theory.

Srivastava, J.N. and Chopra, D.V., "Balanced Arrays of Strength 4, 2 Symbols, and $6 \geq \mu_2 \geq 2$. Submitted for publication.

Yates, F. (1937), "The Design and Analysis of Factorial Experiments", Imperial Bureau of Soil Science, Technical Communication No. 35, Harpenden, England.