

ON THE THEORY OF CONNECTED DESIGNS.

I. CHARACTERIZATION

BU-425-M

J. A. Eccleston and A. Hedayat
Cornell University

July, 1972

Abstract

The concept of connectedness in the theory of block designs is due to Bose. Connectedness is an important property which every block design must possess if it is to provide an unbiased estimator for all elementary treatment contrasts under the usual linear additive model. While Bose has defined this concept in the form of chains between blocks and treatments, Chakrabarti has equivalently defined this concept in terms of the rank of the coefficient matrix or the information matrix of the design.

The notion of connectedness is not in general related to any optimality criteria, i.e., it is quite possible that, for the given $v, b; r_1, r_2, \dots, r_v; k_1, k_2, \dots, k_b$ the parameters of the design, an arbitrary connected design may happen to be the "worst" possible one. This means that one should study and classify the family of connected designs from an optimality point of view. This problem can be tackled in two different ways. (i) Search for the optimal design under the given optimality criterion. (ii) Decompose the family of connected designs into "meaningful" subclasses and study the optimality of each subclass. While approach (i) seems to be natural, it is certainly hard and in some cases impossible, given our present mathematical machineries. Approach (ii) depends heavily on the way one might classify the family of connected designs. An arbitrary partition is certainly useless and will lead us nowhere. We will use the approach (ii) and the following considerations motivated our classifications. We observed that for some connected designs not every observation participates in the least squares estimation of contrasts. This consideration suggested to us the possibility that a connected design which has the property that every observation participates in such an estimation is "better" than one which lacks this property. Thus we classified the family of connected designs into three subclasses: locally connected, globally connected and pseudo-globally connected designs. Basically, a locally connected design is one in which not all the observations participate in the estimation. A globally connected design is one in which all the observations participate in the estimation. Finally, a pseudo-globally connected design is a compromise between locally and globally connected designs.

In this paper we limit ourselves to the characterization of these three types of connected designs. Optimality of these designs will be the subject of paper no. II. The following highlights the content of the present paper. In section 2 the three above mentioned classes of connected designs are rigorously defined. The characterization of each class is dealt with in section 3. Invariance properties, the problem of composing connected designs and the graph theoretic analogy are considered in the final two sections of the paper.

ON THE THEORY OF CONNECTED DESIGNS.

I. CHARACTERIZATION

J. A. Eccleston and A. Hedayat
Cornell University

1. Introduction and Summary. The concept of connectedness in the theory of block designs is due to Bose (1947). Connectedness is an important property which every block design must possess if it is to provide an unbiased estimator for all elementary treatment contrasts under the usual linear additive model. While Bose has defined this concept in the form of chains between blocks and treatments, Chakrabarti (1963) has equivalently defined this concept in terms of the rank of the coefficient matrix or the information matrix of the design.

The notion of connectedness is not in general related to any optimality criteria, i.e., it is quite possible that, for the given $v, b; r_1, r_2, \dots, r_v; k_1, k_2, \dots, k_b$ the parameters of the design, an arbitrary connected design may happen to be the "worst" possible one. This means that one should study and classify the family of connected designs from an optimality point of view. This problem can be tackled in two different ways. (i) Search for the optimal design under the given optimality criterion. (ii) Decompose the family of connected designs into "meaningful" subclasses and study the optimality of each subclass. While approach (i) seems to be natural, it is certainly hard and in some cases impossible, given our present mathematical machineries. Approach (ii) depends heavily on the way one might classify the family of connected designs. An arbitrary partition is certainly useless and will lead us nowhere. We will use the approach (ii) and the following considerations

This research was supported by National Institutes of Health Research Grant No. 5 R01-GM-05900 at Cornell University.

AMS subject classification. 62K05.

Key words. Block design, connected block design, locally connected, pseudo-globally connected, globally connected.

motivated our classifications. We observed that for some connected designs not every observation participates in the least squares estimation of contrasts.

This consideration suggested to us the possibility that a connected design which has the property that every observation participates in such an estimation is "better" than one which lacks this property. Thus we classified the family of connected designs into three subclasses: locally connected, globally connected and pseudo-globally connected designs. Basically, a locally connected design is one in which not all the observations participate in the estimation. A globally connected design is one in which all the observations participate in the estimation. Finally, a pseudo-globally connected design is a compromise between locally and globally connected designs.

In this paper we limit ourselves to the characterization of these three types of connected designs. Optimality of these designs will be the subject of paper no. II. The following highlights the content of the present paper. In section 2 the three above mentioned classes of connected designs are rigorously defined. The characterization of each class is dealt with in section 3. Invariance properties, the problem of composing connected designs and the graph theoretic analogy are considered in the final two sections of the paper.

2. Preliminaries and Definitions. The concept of connected block designs was introduced by Bose (1947). Before presenting Bose's definition let us define a block design. Let $\Omega = \{1, 2, \dots, v\}$ be a set of v treatments assigned to b blocks of size k_u , $u = 1, 2, \dots, b$ and treatment i is replicated r_i times. We denote this general block design by $D = \{B_1, B_2, \dots, B_b\}$, where B_u is the u^{th} block. The statistical analysis of interest in this paper is the intrablock analysis with the model:

$$E(y_{iu}) = \mu + t_i + \beta_u$$

where y_{iu} is the observed response of the i^{th} treatment in the u^{th} block, μ = mean effect, t_i = the effect of treatment i , and β_u = the effect of the u^{th} block.

From the normal equations we have

$$(2.1) \quad \underline{C}\hat{\underline{t}} = \underline{Q}$$

where $\hat{\underline{t}}$ is a solution of (2.1) and called the vector of estimated treatment effects

$$(2.2) \quad \underline{C} = \text{diag}(r_1, r_2, \dots, r_v) - \underline{N} \text{diag}(k_1^{-1}, k_2^{-1}, \dots, k_b^{-1}) \underline{N}'$$

or

$$\underline{C} = \underline{R} - \underline{N}\underline{K}^{-1}\underline{N}' .$$

\underline{N}' is the transpose of \underline{N} , the incidence matrix of the design

$$\underline{Q} = \underline{T} - \underline{N}\underline{K}^{-1}\underline{B}$$

\underline{T} = column vector of treatment totals.

\underline{B} = column vector of block totals.

Equation (2.1) is known as the equation for estimating the treatment effects and the matrix defined by (2.2) is the well known coefficient matrix. Obviously, the C matrix plays a decisive role in the estimation of contrasts and hence the connectedness of designs.

Bose defined connectedness as follows:

"A treatment and block are said to be associated if the treatment is contained in the block. Two treatments, two blocks, or a treatment and a block may be said to be connected if it is possible to pass from one to the other by means of a chain consisting alternately of blocks and treatments such that any two members of a chain are associated. A design (or a portion of a design) is said to be a connected

design (or a connected portion of a design) if every block or treatment of the design (or a portion of the design) is connected to every other."

Unbiased estimators of an elementary treatment contrast can be obtained directly from the chains connecting the treatments of the contrast. For example, consider a block design where block B_1 contains treatments (i, i_1) , block B_2 contains treatments (i_1, i_2) , \dots , block B_h contains treatments (i_{h-1}, i_h) and block B_{h+1} contains treatments (i_h, j) . Then treatments i and j are connected through the chain $iB_1i_1B_2i_2\cdots i_{h-1}B_hi_hB_{h+1}j$ and an unbiased estimator of $t_i - t_j$ is obtained from this chain by the following linear function of the corresponding observations $y_{i1} - y_{i_11} + y_{i_12} - y_{i_22} + \cdots + y_{i_{h-1}h} - y_{i_hh} + y_{i_{h+1}h+1} - y_{jh+1}$. Chains of the form $iB_u i$ are meaningless and should not appear as part of any chain between two treatments. It is interesting to note that if the design is connected with respect to treatments it is also connected with respect to blocks and all elementary contrasts between blocks are estimable, i.e., $\beta_u - \beta_{u'}$ is estimable for all $u, u' = 1, 2, \dots, b$, $u \neq u'$. Chakrabarti (1963) defines a design to be connected if its C matrix has rank $v - 1$, and has proved that his definition of connected designs is equivalent to that of Bose (1947).

Example 2.1. Consider the following design:

$$D: \begin{array}{ccc} B_1 & B_2 & B_3 \\ \begin{array}{|c|} \hline 1 \\ \hline 2 \\ \hline \end{array} & \begin{array}{|c|} \hline 2 \\ \hline 3 \\ \hline \end{array} & \begin{array}{|c|} \hline 3 \\ \hline 4 \\ \hline \end{array} \end{array}, \quad \Omega = \{1, 2, 3\} .$$

D has the following chains:

$$\begin{array}{l} 1B_1^2, \quad 1B_1^2B_2^3, \quad 1B_1^2B_2^3B_3^4 \\ 2B_2^3, \quad 2B_2^3B_3^4, \quad 3B_3^4 . \end{array}$$

Every treatment is connected by a chain to every other treatment, and therefore

the design is connected. The design has C -matrix

$$C = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}$$

and the rank of C is 3 ($= v-1$) and thus, as before, the design is connected. If treatment 2 or 3 of B_2 is deleted the resulting design is not connected.

The original definition of connectedness is extended and generalized to further classify connected designs as either locally, globally or pseudo-globally connected. Locally connected designs are defined the same as the connected designs of Bose (1947) and Chakrabarti (1963). Example 2.1 exhibits a locally connected design. Hedayat (1971) defines two treatments to be globally connected as follows:

Definition 2.1. Two treatments i and j , $i \neq j$, of a block design are said to be globally connected if each replicate of i is connected by a chain, as defined by Bose (1947) to each replicate of j .

Denote the x^{th} replicate of treatment i as i^x .

Example 2.2. Consider the following block design:

	B_1	B_2
D:	1^1	1^2
	2^1	2^2
	3^1	3^2

The chains between the replicates of treatments 1 and 2 are

$$\begin{array}{ll} 1^1 B_1 2^1 & 1^2 B_2 3 B_1 2^1 \\ 1^1 B_1 3 B_2 2^2 & 1^2 B_2 2^2 \end{array} .$$

For treatments 1 and 3

$$\begin{array}{ll} 1^1 B_1 3^1 & 1^2 B_2 3^2 \\ 1^1 B_1 2 B_2 3^2 & 1^2 B_2 2 B_1 3^1 \end{array} .$$

For treatments 2 and 3

$$\begin{array}{ll} 2^1 B_1 3^1 & 2^2 B_2 3^2 \\ 2^1 B_1 1 B_2 3^2 & 2^2 B_2 1 B_1 3^1 \end{array} .$$

Each pair of treatments is globally connected. By the deletion of any treatment from the design the remaining pairs of treatments will not be globally connected. Eccleston (1972) defines pseudo-globally connectedness as follows:

Definition 2.2. Two treatments i and j , $i \neq j$, of a block design are said to be pseudo-globally connected if each replicate of i is connected by a chain, as defined by Bose, to at least one replicate of j and vice versa.

Example 2.3. Consider the following block design:

$$D: \begin{array}{cc} B_1 & B_2 \\ \boxed{\begin{array}{c} 1^1 \\ 2^1 \end{array}} & \boxed{\begin{array}{c} 1^2 \\ 2^2 \end{array}} \end{array} .$$

The design has the following chains: $1^1 B_1 2^1$ and $1^2 B_2 2^2$; therefore, treatments 1 and 2 are pseudo-globally connected. Note that treatments 1 and 2 are not

globally connected.

In the following definition and lemma we use the term "x connected" where x can mean locally, globally or pseudo-globally.

Definition 2.3. A block design is said to be x connected if every pair of treatments is x connected.

Examples 2.1, 2.2, and 2.3 exhibit a locally connected design, a globally connected design and a pseudo-globally connected design, respectively.

If we allow a treatment to be x connected to itself then the relation $R(x)$, treatments i and j are x connected, defines an equivalence relation on Ω . We now have the following lemma.

Lemma 2.1. A design is x connected if and only if under the equivalence relation $R(x)$ there is only one equivalence class.

3. Characterization.

A. Locally Connected Designs. In this subsection several new results for determining whether or not a design is locally connected are given. Some corollaries and rules for special cases are also given along with a few examples which demonstrate the usefulness of these results. First, let us review some results from the literature.

Gateley (1962) and Weeks and Williams (1964) give conditions for a n-way crossed classification design with no interactions to be locally connected. Gateley's theorems involve the rank of the design matrix and for block designs ($n = 2$), it is equivalent to Chakrabarti's rank of C definition. The procedure of Weeks and Williams is too lengthy to present here, and the reader is referred to their 1964 paper or Searle (1971).

One should note that Chakrabarti's 1963 paper contains many important results on the \underline{C} -matrix and is considered a major contribution to the theory of connected designs. Rules from Chakrabarti (1963), Hedayat (1971) and Eccleston (1972) which help in determining for special cases, whether or not a design is locally connected, follow.

(a) D is locally connected if every element in its \underline{C} matrix is different from zero.

(b) D is locally connected if its \underline{C} matrix contains a row (column) of non-zero elements.

(c) D is locally connected if there is at least one non-zero element in row i , $i = 1, 2, \dots, v-1$ above the non-zero elements in the last row of its \underline{C} matrix.

(d) D is locally connected if there are more than $v + t$ non-zero elements in row i , $i = 1, 2, \dots, v-1$ and there are only t non-zero elements in the last row of its \underline{C} matrix.

(e) D is not locally connected if there is any zero on the main diagonal of its \underline{C} matrix.

(f) D is not locally connected if there is an i and j such that in its \underline{C} matrix we have $c_{ii}c_{jj} \leq c_{ij}^2$ and $v > 2$.

(g) D is not locally connected if the largest element (in absolute sense) in its \underline{C} matrix does not lie on its main diagonal and $v > 2$.

(h) D is locally connected if \underline{N} has a row or column with no zero elements, i.e., if a treatment appears in every block or a block contains every treatment, then D is locally connected.

Theorem 3.2. D is locally connected if and only if there exists a set

$D^* = \{B_1^*, B_2^*, \dots, B_b^* \mid B_s^* \in D \forall s=1,2,\dots,b \text{ and there exists a } q < p \text{ such that}$

$$B_p^* \cap B_q^* \neq \emptyset \forall p=2,3,\dots,b\} .$$

Proof:

(i) Sufficiency. The existence of D^* implies that every treatment must appear in a block that contains at least two treatments. Thus each B_s^* must intersect with a B_r^* $r \neq s$, that contains at least two treatments and the union of all blocks containing two treatments contains Ω . Hence we can construct a chain that passes through all the blocks containing two or more treatments and thus pass through every treatment.

(ii) Necessity. If D^* does not exist then there is a B_p^* for which no B_q^* exists such that $B_p^* \cap B_q^* \neq \emptyset$, $q < p$, and the B_s^* 's can be grouped into disjoint sets of B_s^* . Thus the treatments contained in these disjoint sets of B_s^* form subsets of connected treatments and D is not locally connected.

Corollary 3.2. A design is locally connected if and only if there exists a chain between two treatments that contain all the treatments or blocks.

Let us consider the set T_i , which has as elements the blocks that contain treatment i , and denote $J = \{T_1, T_2, \dots, T_v\}$.

Theorem 3.3. D is locally connected if and only if there exists a set

$J^* = \{T_1^*, T_2^*, \dots, T_v^* \mid T_i^* \in J \forall i=1,2,\dots,v \text{ and there exists a } j < i \text{ such that}$

$$T_i^* \cap T_j^* \neq \emptyset \forall i=2,3,\dots,v\} .$$

Proof: This proof is analogous to that of theorem 3.2.

If treatment i and j are connected by a chain we write this as $[ij]$. Define the operator \cdot (dot) by $[ij] \cdot [jk] = [ik]$; i.e., if i and j are connected and j and k are connected then, obviously, i and k are connected by a chain. Also, if i and j are connected by a chain then j and i are connected by a chain; i.e., $[ij] = [ji]$. It should be noted that if a design is locally connected then there are $v(v-1)$ chains, excluding the chains of $[ii]$.

Theorem 3.4. D is locally connected if and only if there is a set, \mathcal{U} , with $v - 1$ elements each of the form $[ij] \in D$, such that under the dot operator, as defined above, the $v(v-1)$ possible chains can be generated.

Before proving the theorem, note that if \mathcal{U} exists every treatment appears in at least one element of \mathcal{U} . Under the dot operator each element gives rise to $v - 1$ other chains plus its reverse; i.e., $[ij] = [ji]$. The total number of chains is $v(v-1)$ since there are $(v-1)(v-2)$ chains by dot operator plus $2(v-1)$ from the elements of \mathcal{U} and their reverses.

Sufficiency part is obvious, and the necessity part follows from the fact that if D is locally connected then every treatment is connected to every other treatment and \mathcal{U} can be easily constructed.

For example 2.1 we have the following sets corresponding to those of theorems 3.2, 3.3, and 3.4, respectively.

$$D^* = \{B_1, B_2, B_3\}, \text{ i.e., } B_1^* = B_1, B_2^* = B_2, \text{ and } B_3^* = B_3$$

$$J^* = \{T_2, T_3, T_1, T_4\}$$

and

$$\mathcal{U} = \{[12], [23], [34]\}.$$

The non-zero elements of $\underline{\underline{NN}}'$ represent the number of chains of the form iB_rj , which is the $[ij]$ element. Thus $(\underline{\underline{NN}}')^2$ is in essence the result of the dot operation between the chains represented by non-zeros in $\underline{\underline{NN}}'$ and in general $(\underline{\underline{NN}}')^a$, $2 \leq a \leq v-1$, is equivalent to the dot operation between the non-zero elements of $(\underline{\underline{NN}}')^{a-1}$ and those of $\underline{\underline{NN}}'$. The longest possible chain between any two treatments is one which contains all the treatments; such a chain could be constructed by the dot operation between $v - 1$ chains of the form iB_rj with distinct B_r 's. Thus the non-zero elements of $(\underline{\underline{NN}}')^{v-1}$ represent those pairs of treatments that are locally connected. We now have the following theorem.

Theorem 3.5. A design is locally connected if and only if its incidence matrix $\underline{\underline{N}}$ has the property that $(\underline{\underline{NN}}')^{v-1}$ has no zero entries.

To demonstrate this theorem consider $\underline{\underline{N}}$ and $(\underline{\underline{NN}}')^3$ of example 2.1. Thus we have

$$\underline{\underline{N}} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}, \quad (\underline{\underline{NN}}')^3 = \begin{bmatrix} 5 & 9 & 5 & 1 \\ 9 & 19 & 15 & 5 \\ 5 & 15 & 19 & 9 \\ 1 & 5 & 9 & 5 \end{bmatrix}$$

As before, the design is locally connected.

Corollary 3.3. In the above theorem $(\underline{\underline{NN}}')^{v-1}$ can be replaced by $(\underline{\underline{N}'\underline{\underline{N}}})^{b-1}$ and the condition remains necessary and sufficient.

Corollary 3.4. If any power of $\underline{\underline{NN}}'$ or $\underline{\underline{N}'\underline{\underline{N}}}$ has a row or column with no zero elements, then the design is locally connected.

Algorithms and further examples for theorems 3.2, 3.3, and 3.4 are given by Eccleston (1972) to demonstrate the practicality of these results.

B. Globally Connected Designs. An advantage of globally connected designs is that when estimating the elementary contrast between the effects of treatments i and j every replicate participates to a maximum yielding $r_i \times r_j$ estimates of $t_i - t_j$ or $t_j - t_i$. As has been shown by Eccleston and Hedayat (1972), the class of connected designs under certain restrictions and constraints contains the optimum design. The following theorem characterizes globally connected designs.

Theorem 3.6. A design D is globally connected if and only if the following conditions hold simultaneously:

- (1) D is locally connected.
- (2) Every block of D contains at least two treatments that occur in more than one block; i.e., for all $B_s \in D$ there exists an i and $j \in B_s$ such that $i \in B_r$ and $j \in B_u$, $u \neq s$ and $r \neq s$.
- (3) For all B_s which there exist i and j such that the treatments belonging to $B_s - \{i, j\}$ do not occur elsewhere, then each of i and j occur in two other blocks.
- (4) Any treatment, i say, that appears in two or more blocks (but not all blocks) must do so in blocks that contain
 - (i) a treatment that appears in two blocks containing i , and two not containing i . That is, $i \in B_r$ and B_s and there exists a $j \in B_r, B_s, B_m$, and B_n where $i \notin B_m$ and $i \notin B_n$,

or

- (ii) two treatments each appearing in a block containing i, and a block not containing i. That is, i and j \in B_r, i and k \in B_s, then j \in B_m and k \in B_n with i \notin B_m and i \notin B_n.

Some of these conditions may seem redundant; however, with a few simple examples it can be shown that this is not the case, see Eccleston (1972). In the following proof by a singleton we mean a block containing exactly one treatment.

Proof of Theorem 3.6.

(i) Sufficiency: Consider any replicate of any treatment, say replicate x of treatment i , and denote as i^x . Then given that the conditions hold, can i^x be connected by a chain to any replicate of any other treatment, say m^y ? Now by condition (2), if $i^x \in B_s$ then there exists a $j \in B_s$ such that we have $i^x B_s j$. Since the design is locally connected we can construct a chain between j and m . If j is connected to m^y , then we are finished. However, if j is connected to m^z , $z \neq y$, then since the blocks containing m^z and m^y satisfy the conditions (2), (3), and (4), a chain between m^z and m^y can be constructed.

(ii) Necessity. (i) Condition (1) is obvious. (ii) If condition (2) is violated then D has a singleton. The treatment belonging to the singleton cannot be connected by a chain to any other treatment and so it follows that D is not globally connected. (iii) If condition (3) is violated by i but not j of block B_s then i occurs in only one other block, B_r say. A chain between $j \in B_s$ and $i \in B_r$ cannot be constructed; consequently, the design is not globally connected. (iv) If condition (4) does not hold for treatment i say, then there is a treatment j which occurs in at least two blocks containing i , and exactly one not containing i , say B_r , or vice versa. It follows that one cannot construct chains between all the replications of j and i , namely the replicate of $j \in B_r$ and any replicate of i . Thus D is not globally connected.

Corollary 3.5. If the same two treatments appear in every block, then the design is globally connected. (The design must have at least three blocks.)

Corollary 3.6. If N has no zero elements, then D is globally connected. (If N has no zero elements, then NN' and $N'N$ have no zero elements.)

Example 3.1. Consider the design

	B_1	B_2	B_3	B_4	B_5
D:	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 4 \\ 3 \end{bmatrix}$	$\begin{bmatrix} 2 \\ 3 \\ 5 \end{bmatrix}$	$\begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}$

By inspecting D it is clear that the design satisfies theorem 3.6.

C. Pseudo-Globally Connected Designs. A pseudo-globally connected design assures one that in estimating elementary contrasts each replicate of the treatments involved is utilized. When estimating elementary treatment contrasts, globally connected designs maximize the use of all replicates of the treatments whereas pseudo-globally connected designs guarantee that no replicates are "wasted". That is, every replicate of each treatment in the contrast is involved at least once in the estimation. As mentioned before, this class of connected designs, under certain conditions, contains the optimum connected design. This is discussed in detail in Eccleston and Hedayat (1972). The following theorem characterizes pseudo-globally connected designs.

Theorem 3.7. A design D is pseudo-globally connected if and only if conditions (1), (2) and (4) of theorem 3.6 hold simultaneously.

The proof is analogous to that of theorem 3.6.

Example 3.2. Consider BIBD(3,3,2,2,1)

D:	B_1	B_2	B_3
	1 ¹	1 ²	2 ²
	2 ¹	3 ¹	3 ²

D satisfies theorem 3.7 and so is pseudo-globally connected.

Corollary 3.7. If a design D is locally connected and each replicate of treatment i is connected by a chain to every other replicate of i, for all $i \in \Omega$ then D is pseudo-globally connected. [Note: If, in addition to the above, condition (3) of theorem 3.6 holds then D is globally connected.]

Further corollaries rules and examples are given by Eccleston (1972).

4. Invariance Properties and the Composition of Connected Designs.

A. Invariance Properties. If a design D on Ω is locally (globally) connected then any of the following can occur and D will remain locally (globally) connected.

(a) For D locally connected: Any new block can be added to D so long as its elements belong to Ω .

(b) For D globally connected:

- (i) any treatment belonging to Ω can be added to any block of D.
- (ii) any new treatment(s) can be added to any block of D.
- (iii) any block belonging to D can be repeated any number of times.
- (iv) if a treatment appears in a block it can be replicated any number of times within that block.

Recall that if a design is globally connected then it is pseudo-globally connected, which also implies that the design is locally connected. Thus the facts in (b) above apply to pseudo-globally and also locally connected designs.

B. The Composition of Connected Designs. Let us consider the proposition of composing two designs that are locally, globally or pseudo-globally connected.

(a) Compositions that yield locally connected designs:

(i) If D_1 and D_2 are locally connected designs on the sets of treatments Ω_1 and Ω_2 , respectively, and $\Omega_1 \cap \Omega_2 = \emptyset$, then the design $\bar{D}_l = D_1 \cup D_2 \cup B$ is locally connected, where B is a block containing at least two treatments, i and j say, such that $i \in \Omega_1$ and $j \in \Omega_2$. The block B forms the link between the two designs D_1 and D_2 . Since i is connected to all treatments in Ω_1 and j to all in Ω_2 then the chain iBj locally connects every pair of treatments of $\Omega_1 \cup \Omega_2$.

(ii) Let D_1 and D_2 be locally connected designs on Ω_1 and Ω_2 , respectively, and if $\Omega_1 \cap \Omega_2 \neq \emptyset$, i.e., Ω_1 and Ω_2 have at least one element in common, then $D_1 \cup D_2$ is a locally connected design.

(b) Compositions that yield globally connected designs.

(i) Consider D_1 and D_2 to be globally connected designs on treatment sets Ω_1 and Ω_2 , respectively, $\Omega_1 \cap \Omega_2 = \emptyset$. As before, $\bar{D}_g = D_1 \cup D_2 \cup B$ where B as above, is locally connected. However, if B contains four treatments $i, j, k,$ and l such that i and $j \in \Omega_1$ and k and $l \in \Omega_2$, also i and j each appear in at least two blocks of D_1 and similarly k and l in D_2 , then \bar{D}_g

is globally connected. Moreover, if B contains three treatments of Ω_1 and three of Ω_2 then \bar{D}_g is globally connected. It is easily shown that \bar{D}_g , with the above B's, satisfies theorem 3.6.

(ii) For $D_1 \cup D_2$ to be globally connected, it is sufficient for D_1 and D_2 each to be globally connected and one of the following:

- (1) $\Omega_1 \cap \Omega_2 = \{i\}$ and i appears in two blocks of D_1 and two of D_2 .
- (2) $\Omega_1 \cap \Omega_2 = \{ij\}$ and i appears in at least one block of D_1 and two of D_2 , while j appears in at least one block of D_2 and two of D_1 .
- (3) $\Omega_1 \cap \Omega_2 = \{i,j,k\}$.

(c) Compositions that yield pseudo-globally connected designs.

(i) Suppose D_1 and D_2 are pseudo-globally connected designs on treatment sets Ω_1 and Ω_2 , respectively and $\Omega_1 \cap \Omega_2 = \emptyset$. As above $\bar{D}_{pg} = D_1 \cup D_2 \cup B$, where B is as in (a), locally connected. However, if i and j belong to two blocks of D_1 and D_2 , respectively, then \bar{D}_{pg} is pseudo-globally connected. Moreover, if B contains 3 treatments i, j and m where i and $j \in \Omega$, and m belongs to two or more blocks of D_2 , then \bar{D}_{pg} is pseudo-globally connected.

(ii) For $D_1 \cup D_2$ to be pseudo-globally connected, it is sufficient that D_1 and D_2 each be pseudo-globally connected and one of the following:

(1) $\Omega_1 \cap \Omega_2 = \{i\}$ and i occurs in two blocks of D_1 and two of D_2 .

(2) $\Omega_1 \cap \Omega_2 = \{i, j\}$.

It is interesting to note that two designs, D_1 and D_2 , can each be not locally connected but their union $D_1 \cup D_2$ may be locally connected. This is obvious since given a locally connected design, D , one can often partition D into locally disconnected subsets. A similar remark is true for globally and pseudo-globally connected designs. The composition of more than two designs would follow along the lines of the above methods but be somewhat more complex.

5. Graph Theoretical Analogy to Connected Designs. A graph G is a mathematical system consisting of two sets V and E . V is a finite nonempty set of p vertices and E is a prescribed set of q unordered pairs of distinct vertices of V . Each pair $e = \{u, v\}$ of vertices in E is an edge of G and e is said to join u and v . We write $e = uv$ and say that u and v are adjacent vertices, vertex u and edge e are incident with each other, as are v and e . Two distinct edges incident with a common vertex are said to be adjacent edges.

A walk of a graph is an alternating sequence of vertices and edges beginning and ending with vertices in which each line is incident with the two vertices immediately preceding and following it. A trail is a walk with all edges distinct and a path is one with all vertices distinct. Harary (1969) defines a graph to be connected if every pair of vertices are joined by a path.

We define the treatments of a design to be the vertices of a graph G , and two vertices are incident if the two treatments belong to the same block. A walk of a graph is equivalent to a treatment-block chain as defined by Bose (1947). Thus, if every pair of vertices is connected by a walk then the design

will be locally connected, and vice versa. Also we can define a design to be locally connected if and only if the graph G is connected.

Analogue to Theorem 3.1. The design D is locally connected if and only if the graph G , as defined above, has only one connected component.

Define the graph $G(D)$ to have as vertices the blocks of the design D and two vertices as incident if the two blocks have at least one treatment in common.

Analogue to Theorem 3.2. D is locally connected if and only if every pair of vertices of $G(D)$ is connected by a walk.

If the T_i , as defined in section 3, are the vertices of graph $G(\mathcal{T})$ and T_i and T_j , $i \neq j$, are incident if there is a $B_s \in T_i$ and a $B_r \in T_j$ such that $B_r \cap B_s \neq \emptyset$, then we have the following:

Analogue to Theorem 3.3. D is locally connected if and only if every pair of vertices of $G(\mathcal{T})$ is connected by a walk.

As before, we can develop some simple rules, in graph theory terms, for determining the local connectedness of D .

- (i) D is not locally connected if any of the above graphs has an isolation vertex.
- (ii) If any vertex, v , of the above graphs has degree (number of edges incident with v) $p - 1$, where p is the number of vertices, then D is locally connected.

The removal or loss of treatments from a design obviously can affect the local connectedness of that design. Knowledge of treatments, which by their removal or loss cause the design to be not locally connected, would usually be

of interest to the experimenter. A similar situation arises in graph theory. Busacker and Saaty (1965) define a vertex v to be a point of articulation of a connected graph if the graph obtained by deleting v and all edges incident with v is disconnected. A graph is said to be separable if it has at least one articulation point.

Lemma 5.1. A necessary and sufficient condition for a vertex v to be a point of articulation is that v lie on all the paths connecting some pair of vertices.

Proof: See Busacker and Saaty (1965).

A matrix of interest in graph theory is the vertex or adjacency matrix V . The element in the (i,j) position of V is the number of edges incident with both vertex i and vertex j . From Busacker and Saaty (1965) we have the following theorem:

Theorem 5.1. The matrix V^n gives the number of walks of length n between any two vertices, where the length of a walk is the number of edges between the beginning and terminating vertices.

The analogous experimental design theory for the above terminology and theory is obvious. A treatment or block is said to be a point of articulation if the design obtained by deleting that treatment or block is not locally connected,

Lemma 5.2. A necessary and sufficient condition for a treatment or block to be a point of articulation is that it lie on all chains connecting some pair of treatments.

If we define the length of a chain to be the number of distinct blocks that appear in the chain, then a treatment matrix can be defined similar to the vertex

matrix of a graph. A treatment matrix Λ has as its (i, j) element the number of blocks that contain both treatments i and j .

Theorem 5.2. The matrix Λ^n gives the number of chains of length n between any two treatments of a design.

Globally and pseudo-globally connected designs were not considered in graph theory terms and it is doubtful if an analogy to theorems 3.6 and 3.7 would be of any use.

6. Acknowledgements. The authors wish to thank Professors J. Kiefer and S. R. Searle for reading the original manuscript and their helpful suggestions.

REFERENCES

- Bose, R. C. (1947). Presidential Address. Proceedings of the 34th Indian Science Congress.
- Busacker, R. and Saaty, T. (1965). Finite Graphs and Networks. McGraw-Hill, New York.
- Chakrabarti, M. C. (1963). On the C-matrix design of experiments. J.I.S.A. 1 8-23.
- Eccleston, J. A. (1972). On the theory of connected designs. Ph.D. Thesis, Cornell University.
- Eccleston, J. A. and Hedayat, A. (1972). On the theory of connected designs. II. Optimality. BU-427-M, Cornell University.
- Gateley, W. Y. (1962). Application of the generalized inverse concept to the theory of linear statistical models. (Unpublished Ph.D. dissertation, Oklahoma State University.)

Harary, F. (1969). Graph Theory. Addison-Wesley, Reading, Mass.

Hedayat, A. (1971). A set of conditions which every locally connected design must satisfy. BU-392-M, Cornell University.

Searle, S. R. (1971). Linear Models. Wiley, New York.

Weeks, D. L. and Williams, D. R. (1964). A note on the determination of connectedness in an N-way cross classification. Technometrics 6 319-324.