Some computational aspects of linear classification models

Abel Geber Mexas
Iowa State University

Follow this and additional works at: https://lib.dr.iastate.edu/rtd
Part of the Statistics and Probability Commons

Recommended Citation
https://lib.dr.iastate.edu/rtd/4185

This Dissertation is brought to you for free and open access by the Iowa State University Capstones, Theses and Dissertations at Iowa State University Digital Repository. It has been accepted for inclusion in Retrospective Theses and Dissertations by an authorized administrator of Iowa State University Digital Repository. For more information, please contact digirep@iastate.edu.
MEXAS, Abel Geber, 1940-
SOME COMPUTATIONAL ASPECTS OF LINEAR CLASSIFICATION MODELS.

Iowa State University, Ph.D., 1970
Statistics

University Microfilms, A XEROX Company, Ann Arbor, Michigan
SOME COMPUTATIONAL ASPECTS OF LINEAR CLASSIFICATION MODELS

by

Abel Geber Mexas

A Dissertation Submitted to the
Graduate Faculty in Partial Fulfillment of
The Requirements for the Degree of
DOCTOR OF PHILOSOPHY

Major Subject: Statistics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

Head of Major Department

Signature was redacted for privacy.

Dean of Graduate College

Iowa State University
Of Science and Technology
Ames, Iowa

1970
# TABLE OF CONTENTS

**I. INTRODUCTION**

A. General Introduction and a Brief Review of Literature 1

B. Scope and Content of the Chapters 5

**II. ANALYSIS OF VARIANCE OF BALANCED COMPLETE EXPERIMENTAL DATA ON A DIGITAL COMPUTER** 13

A. General 13

B. Definitions 15

C. Structures 18

D. Structure Matrices 22

E. Lattice Properties 24

F. Admissible Means 33

G. Components of Variation 41

H. Expected Mean Squares 45

I. Extension to Balanced Incomplete Samples from Balanced Complete Populations 57

J. A Comparison of the Development 59

**III. THE NUMBER OF ADMISSIBLE SETS IN AN ARBITRARY STRUCTURE, AND THE NUMBER OF STRUCTURES CONTAINING AN ARBITRARY NUMBER OF FACTORS** 66

A. Two Problems 66
B. Notational Conventions

C. Number of Admissible Sets in an
   Arbitrary Structure

D. On the Number of Structures Containing
   an Arbitrary Number of Factors
   1. Number of Structures with Two Factors
      not Directly Nested in the Mean 84
   2. Number of Structures with Three Factors
      not Directly Nested in the Mean 97
   3. Extension of the Induction Procedure 128

IV. IDENTIFICATION OF ESTIMABLE FUNCTIONS IN
    ARBITRARY CLASSIFICATION ARRANGEMENTS AND
    NON-ORTHOGONAL ANALYSIS OF VARIANCE 131
    A. The General Problem 131
    B. A General Method of Computing the
       Non-orthogonal Analysis of Variance 136
    C. Bradley's Algorithm 148
    D. Methods Using Conditional Inverses 153
    E. Wilkinson's Recursive Algorithm for
       A.O.V. 159
    F. Contribution of Elston and Bush 168

V. ESTIMATION AND ITS RELATIONSHIP TO
    CONNECTEDNESS, REDUCIBILITY AND GRAPHS 170
    A. Introduction 170
    B. Bose's Theorem and its Extension 171
C. The Equivalence Between Connectedness,
   Irreducibility of the Incidence Matrix,
   and a Property of its Graph  178
D. The Three Way Additive Classification  184

VI. ACKNOWLEDGEMENTS  200

VII. BIBLIOGRAPHY  201

VIII. APPENDIX  204
       Computer Program for Counting the Number
       of Structures, q = 3.  204
I. INTRODUCTION

A. General Introduction and a Brief Review of Literature

This work studies various problems in the context of a general linear classification model. This chapter defines the problems, the assumptions in each problem, and the extent of the solutions contained herein. All the problems considered involve digital computation in one way or another. Some of the problems were prompted by existing needs for better and more general computer programs. Other problems of general theoretic interest were investigated with the objective of defining general computational algorithms. In still other problems the computer was found necessary in evaluating expressions attained through logical or other theoretical considerations.

The writer's interest in this topic was stimulated by three years experience in statistical computing at the Iowa State University Numerical Analysis Section of the Statistical Laboratory. During this time consultants relied heavily on an analysis of variance program called AARDVARK (20), designed to permit machine computations directed by algebraic specifications of the statistical problem. This program evolved from an algorithm for the analysis of variance published in 1956 by H.O. Hartley (13), which stimulated activity and interest in computing. W.J. Hemmerle (16) detailed the algorithm and the language through which the Hartley techniques were implemented in a manner useful to a wide class of users of analysis of variance. The AARDVARK program
has been used extensively since 1963, undergoing many changes in available options and even basic computer hardware for which the program would be applicable. Similar and parallel efforts at implementation of general purpose statistical systems have been accomplished in recent years.

Yates (33), (34), in 1962, gave impetus to the need for better and new statistical computing systems, and in 1966, Yates and Anderson (35) published a general computer program for the analysis of factorial experiments. Nelder (21), in 1964, sketched a general computer program to handle the analysis of experiments with orthogonal block structure, to which treatments are applied randomly and under the assumptions of unit-treatment activity. Bock (1) published a detailed programming system for univariate and multivariate analysis of variance. Dixon (7) was the editor of a series of statistical programs developed at the University of California (Los Angeles) known as BIO-MED or BMD programs, which have been widely distributed and used in data analysis.

The systems implemented offer varying degrees of flexibility and output options. One primary difference between the existing programs is the basic philosophy under which the systems were designed. AAROWARK was aimed at simplicity of problem specification, but while it accomplished this extremely well, the system suffered a considerable loss of flexibility for modification and dependence on the particular computer configuration available at the time of its development. Bock's programs are highly geared to efficient computation, but sacrifice simplicity of use. For example, in the analysis of variance of incomplete classification data, the user is required to input a basis for the vector.
space of estimable functions. The BIO-MED programs are a compromise between simplicity of use and efficient algorithms. However, they are in no way connected as a system, but rather are a collection of individual programs each having limited and specific capabilities. Experiences with the use and maintenance of some of these systems revealed that a major difficulty in effecting accurate and prompt modifications stemmed from obscurities in the intent and purpose of certain intricate computations. Hence, a study of the underlying pattern of computations was undertaken with a view of reducing the description of the operations and the programming of these to the most rudimentary form possible.

Boolean operations on binary controlling arrays had been suggested by Carney (4), (5), and this motivated the development described in Chapter II. The guiding theoretical principles are those formulated by Kempthorne (18), (19), Wilk (27), (28), Wilk and Kempthorne (29), (30), (31), Zyskind (36), (37), (38), Zyskind et.al. (39), Throckmorton (24), (19), and White (26), (19), (39).

The study in Chapter II of the lattice properties of experimental structures led to the investigation of techniques useful in enumerating the structures that are possible with N factors, where two structures are considered distinct if the nesting relationships between the factors are different apart from a permutation of factor names. This problem was suggested by Kempthorne in the work of Throckmorton (24) and Gilbert (12). Throckmorton used Hasse diagrams to represent the structures containing five factors or less. Gilbert (12) studied the enumeration problem for six factors. If N is the number of factors and P(N) the number of structures, then their results provide the following numbers:
Neither author attempted a formal and general enumeration of the type described in Chapter III.

The remainder of the work is devoted to arbitrary experimental arrangements, or equivalently, the analysis of classification data that may contain unequal and disproportionate subclass numbers, with an arbitrary number of the subclass numbers equal to zero. The problem posed is the identification of functions estimable with this type of data.

Chapter IV exhibits the problem of developing general programs for the analysis of classification arrangements arbitrarily not of maximal rank. The chapter reviews critically some of the well known methods of analysis which seem relevant to this class of situations, specifically, the technique for non-orthogonal analysis contained in AARDVARK and the recently proposed methods of Bradley (3), Wilkinson (32), and Elston and Bush (8). Additionally, some techniques which utilize conditional inverses are briefly discussed. Chapter V explores conditions equivalent to maximality rank in a two-way classification. Concepts of connectedness, as defined by Bose (2), irreducibility of matrices, and strong connectedness of a graph associated with matrices are shown to be related to estimability. The concepts lead to the development of a scheme useful in determining the linear parametric functions, within a factor of classification, that can be estimated from a data set of arbitrary incidence.
B. Scope and Content of the Chapters

Chapter II is devoted entirely to a study of logical sequences of computations leading to the analysis of variance. The objective is an algorithm immediately programmable in FORTRAN, PL-1, or APL. The structures considered are generally assumed balanced and complete as defined in the chapter. Otherwise the structures may have arbitrary nesting and crossing relationships among the factors. The primary result is a delineation of strictly logical or Boolean operators that define and control all computations, thus simplifying the otherwise complex descriptions of sequences of operations applicable to each of various subclasses of the class of structures considered.

Section A of Chapter II provides a brief bibliography of the work in the general area of classification models and analysis of variance techniques on which the chapter is based. A few brief definitions of terms frequently used in the chapter are given in Section B.

Section C exhibits various existing representations of data structures. Section D provides a new and basic representation of a structure in terms of binary numbers capable of being stored and manipulated in the memory of a digital computer. This representation is termed the structure matrix. Its construction is defined in terms of simple logical "and" and "or" operations.

Section E explores the advantages of the binary matrix representation of a structure. Primarily, the algebraic properties of structures are explored in a general way, leading to the demonstration that
structures form a distributive lattice. A weaker class of lattices, namely, modular lattices, possess all the required characteristics used implicitly in later algorithms. However, the study of algebraic properties of structures was pursued slightly further as interesting in itself. This culminated in the realization that structure lattices lack the properties of complemented lattices, i.e., if \( a \in J \), where \( J \) is an arbitrary structure, there does not exist \( a' \) such that \( a \lor a' = 1 \), and \( a \land a' = 0 \), where \( \lor \) denotes the least upper bound (l.u.b.), and \( \land \) the greatest lower bound (g.l.b.). This property of a complemented lattice is necessary for the class of lattices defined by experimental structures to possess the richer properties of a Boolean algebra. The structure matrix, through a sequence of logical operations on its rows, leads to a binary matrix termed the full structure matrix. This provides a binary representation of all admissible terms in a classification model. The lattice properties of the full structure matrix are seen to hold when the l.u.b. and g.l.b. operators are taken to be the Boolean sum and Boolean product, respectively.

More significant, computationally, Section \( \text{F} \) defines logical operations which control the computation of all admissible means for such structures. These operations, applied to the rows of the full structure matrix, control the sequence by which means for some model terms would be formed from other strings of admissible partial means. The intent and conclusion is an algorithm efficient in minimizing arithmetic operations and core storage utilization. The latter problem, conserving memory, is considered paramount in describing any general
purpose algorithm because flexibility and applicability are greatly reduced when severe size limitations are imposed. The algorithm, as defined in this section, requires only sufficient array allocation to store twice the number of observations. By comparison, AARDVAR (20), as presently implemented, requires all data, admissible partial means, and components to be stored in core simultaneously.

In Section G, logical operations on the rows of the full structure matrix and an associated matrix termed the rightmost bracket matrix define the logical construction of components from the strings of means. For completeness, the construction of a matrix giving the inverse relation of means defined as linear combinations of the components is also presented in terms of logical operations. Section H translates into logical operations the definitions of quantities called cap sigmas which are linear functions of the components of variation.

A well known theorem giving the expected mean squares (EMS) for each source of variation as a linear function of the cap sigmas is also translated into simply programmed logical operations. The section concludes by providing the modifications that arise in the EMS's when factors are fixed or random, the more general case being when all factors are considered semi-random.

Section I calls attention to possible extensions of the preceding logical development to a wider class of structures than the balanced complete class by using the results of White (26), who gave relatively simple expressions for the EMS in situations characterized by balanced complete populations. Section J gives a formal comparison of the
computational methods suggested in the chapter to those proposed by Nelder (21).

Two problems in connection with $N$-factor structures with arbitrary nesting relationships among the factors are considered in Chapter III. The first and simpler problem is the attainment of a programmable expression, giving the number of admissible sets corresponding to that structure. The procedural logic leading to the resulting expression, as given in Section C, was invaluable in providing a fruitful method of investigating the second and more complex problem of determining the number of distinct structures containing $N$ factors. With the results of Chapter II, this, the principal problem of the third chapter, can be equivalently expressed in general algebraic terms. Definitions 2.16 and 2.17 define, respectively, a partially ordered set and a lattice. Theorem 2.1 states that a factorial structure has the properties of a distributive lattice. In order to clarify what is meant by distinct structures the following definition is necessary.

**Definition 1.1**

Let $S_1 = \{\{a_i\}, \leq_1\}$ and $S_2 = \{\{b_j\}, \leq_2\}$ be two lattices. A 1-1 mapping $a_i \rightarrow b_j$ of a lattice $S_1$ onto a lattice $S_2$ is an isomorphism if and only if $a_i \leq a_i'$ in $S_1$ implies, and is implied, by $b_j \leq b_j'$ in $S_2$.

Thus, two structures are considered distinct if and only if one is not an isomorphism of the other. For example, the set of integers $S = \{1, 3, 4, 12\}$ forms a lattice with respect to the partial ordering relation defined by: if $a, b \in S$, then $a \leq b$ if there exists $k, a$
positive integer, such that \( b = ka \).

Similarly, the power set of the set \( A = \{\emptyset, \{\emptyset\}\} \), where \( \emptyset \) denotes the empty set, is given by \( S_2 = \{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset\}, \{\emptyset\}, \{\emptyset\}, \{\emptyset\}, \{\emptyset\}, \{\emptyset\} \}. \) \( S_2 \) forms a lattice with respect to the partial ordering relation defined by the rule: if \( A, B \in S_2 \), then \( A \leq B \) if \( A \) is a subset of \( B \). However, the two lattices are isomorphic with respect to the 1-1 mapping defined by \( 12 \rightarrow \{\emptyset, \{\emptyset\}\}; \quad 4 \rightarrow \{\{\emptyset\}\}; \quad 3 \rightarrow \{\emptyset\} \) and \( \emptyset \rightarrow 1 \). The two lattices have identical structure since they have corresponding ordering relationships among their respective numbers. Diagrammatically, one represents the elements of the lattice by points, and connects each pair of elements \( a, b \), satisfying \( a \geq b \), by a descending broken line from \( a \) to \( b \). In the preceding, the resulting structure diagrams are:

\[
\begin{array}{ccc}
12 & 3 & 4 \\
\downarrow & \downarrow & \downarrow \\
1 & {\{\emptyset, \{\emptyset\}\}} & \{\{\emptyset\}\} \\
\end{array}
\]

The problem considered in Section D is the enumeration of non-isomorphic lattices of \( N + 2 \) elements. It can be easily seen that every lattice contains a unique element which is an upper bound of every other element in the lattice, and a unique element which is a lower bound of every other element in the lattice. These two elements are called the \( 1, 0 \) elements of the lattice, respectively, and can be made to correspond to the mean and error term in experimental structures. In a lattice \( S \), if \( a, b \), are elements of \( S \), \( a \geq b \), and there
does not exist an element \( c \) in \( S \), such that \( a \succ c \succ b \), then \( a \) is said to be a \textit{cover} of \( b \). In experimental structures, for the same condition, the factor corresponding to \( a \) is said to be a direct nestor of the factor corresponding to \( b \). The approach to the enumeration problem as developed in Section D is an induction with respect to \( q \), where \( q \) is the number of nested factors in the lattice having a cover distinct from the 1-element.

The principal results of Chapter III include a characterization of lattices, with an arbitrary number of elements for the cases \( q=0, 1, 2, 3 \), in terms of 0, 1, 3, 8, -tuples, respectively. A polynomial expression in \( N + 2 \), the number of elements of the lattice, is obtained empirically for \( q' = 2 \). The general induction step for arbitrary \( q \) remains an unsolved problem and is subject to further investigations.

Chapter IV is in part devoted to the review of existing computer methodology applicable in the analysis of classification data of arbitrary incidence. The techniques which are known to be useful, when all the cells are filled with unequal and disproportionate numbers of observations, are examined as a starting point for the development. For this class of situations, the algorithms, which are generally applied, take particular advantage of the a priori knowledge of estimable and non-estimable parametric functions. In dealing with arbitrarily arranged classification data it is not generally known what is estimable. While it is shown that for situations of maximal rank, a full rank reparametrization can be obtained efficiently without construction of the coefficient matrix of the design model; that the
procedure can not be easily modified for the case of unobserved cells is made manifest. For certain algorithms, e.g. AARDVARK non-orthogonal analysis of variance, the results are shown by counterexamples to be incorrect in certain situations. In the opinion of the author, the analysis of classification data of arbitrary incidence requires the attainment of a particular type of basis for the row space of the design coefficient matrix. The basis of interest describes what can be estimated from the given data, and should exhibit explicitly all the independent estimable functions within each factor of classification.

Bradley (3) proposed a scheme for identifying a basis for the space of estimable functions in the case of a totally arbitrary experimental arrangement. This procedure is discussed and shown not to provide, generally, the estimability information required by most data analysts. The use of a sequence conditional inverses is shown to lead to the desired estimable functions, however, this technique is computationally discounted because of the great number of operations required for its general use. A brief exposition of a recursive program for analysis of variance developed by Wilkinson (32) is presented, and shown to be generally non-applicable to the most general class of classificatory situation considered in the Chapter. A procedure based on knowledge of the sum of the rows of the design matrix is explored, and the class of situations to which this procedure applies is defined.

The problem of identifying functions that may be estimated unbiasedly is significantly different in additive models from that in models containing interactive terms. In the latter case, Elston and Bush (8) have exhibited a set of hypotheses that are testable with the arbitrary crossed
classification model. The chapter contains a critical evaluation of the authors' conclusions in so far as they pertain to the general identification problem previously mentioned.

Chapter V establishes and exploits the equivalences between the concept of estimability of all contrasts in parameters of single factors of classification, connectedness of a design as a criterion equivalent to maximal rank, irreducibility of certain incidence matrices, and directed graphs associated with a matrix. It is the strong opinion of the author that these relationships can be fruitfully employed in diverse ways to provide efficient, although somewhat nonstandard, computational techniques. In particular, the present work describes a use of the concepts to identify estimable main effect contrasts from the indices associated with each observation. The cases considered explicitly are restricted to additive models with fewer than four factors.

The concept of connectedness in the context of a block by treatment additive model was introduced by Bose (2) as early as 1947. Scheffe (23) in a footnote, refers to the Bose criterion as relevant in defining a condition on the pattern of non-empty cells of a two way layout for the estimability of contrasts of interest. Elston and Bush (8) used the criterion to provide hypotheses testable in a two-way interactive model with missing cells. The authors concluded that it is possible to derive a reasonable test of hypotheses for the main effects whatever the pattern of empty subclasses, provided only that the filled classes form a connected design. It is also stated that the principles the authors employ computationally in examples based on two way classifications are exactly the same as those applicable to higher-way classifications. It would
appear that this is not the case. Chapter V exhibits the difficulties of applying the connectedness criterion of Bose to models containing more than two factors. The strong connectedness property of a graph is very relevant and is exhibited as being more generally applicable than connectedness.
II. ANALYSIS OF VARIANCE OF BALANCED COMPLETE EXPERIMENTAL DATA ON A DIGITAL COMPUTER

A. General

The applications of linear model theory to the analysis of structure data has been given new impetus by recent computer advances such as direct accessing methods, peripheral storage devices, remote terminals, and new high-level languages like APL and PL-I. As a direct result the quantities of data and the size of the data sets to which standard linear model theory is being applied have increased enormously, and to date seemingly without bound. As a result, existing programs are limited in size or scope and are difficult to modify. This chapter develops practical methods for the computation of the essential elements in the analysis of data possessing sufficient structure to be classified as a balanced and complete. The results include an exploration of the lattice properties of such structures, and an algorithm highly adapted to the advanced state of computer technology.

A development of the analysis of the general experiment was formulated by Kempthorne (18), Kempthorne, et al. (19), Wilk (27), (28), Wilk and Kempthorne (29), (30), (31), Zyskind (36), (37), (38), Zyskind, et al. (39), Throckmorton (24), and in (19), and White (26), and in (19), and (39). This development results in an axiomatic decomposition of the total sums of squares into components. These authors emphasized the "derived linear" rather than "assumed linear" model. Briefly stated, the distinction is that the analysis requires no assumptions concerning the form of the
response as a function of levels of factors. Instead, the analysis proceeds from a more basic assumption that the response is a function of experimental material effect and treatment effects. The two effects may be additive or not additive. Each effect is decomposed by a prescribed identity relation and under a valid randomization the variance-covariance matrix of the observational vector is deduced. Because of the randomization, it is possible to view the realization of the experiment as a random sample from a conceptual population of possible outcomes. This leads to the concept of a "population structure". The sample obtained from this conceptual population has an "observational structure" which may or may not be identical to the structure of the population. The combination of these two structures is termed an "experimental structure", and has the property of containing necessary and sufficient information for possible inferences.

The work here described is essentially concerned with computational consequences derived from this theoretical development. Thus, it is at first assumed that the data have balanced complete structure and that the experimenter can describe the nesting relationships to the computer. The algorithm provides convenient methods to convey and represent the structure in a digital computer and details subsequent computational processes through simple meaningful and logical operators resulting in formal computational methods of obvious advantage.
B. Definitions

For purposes of ready reference, clarity of scope, and intended usage, some frequently occurring terms are here defined. The definitions have been previously given by White in (26), and in Zyskind, et. al. (39), Throckmorton (24), and in Kempthorne et al. (19), and Zyskind (36), (37), (39).

Suppose there exist data sets $Y_i$, such that each corresponds to observed or conceptual responses on some variable. The structure of each data set is imposed by a set of factors $Z = \{A, B, C, \ldots\}$.

**Definition 2.1**

A factor is a partition of a data set $Y_i$ into disjoint non-empty subsets.

**Definition 2.2**

The subsets of the data set, defined by any factor, are called the levels of the factor. For a factor $A$, the levels may be denoted $A = \{a_1, a_2, \ldots\}$.

**Definition 2.3**

A choice of one level from each factor of a set of factors $Z = \{A, B, C, \ldots\}$ is called a combination of the set of factors, and is denoted by $z = \{a_i, b_j, c_k, \ldots\}$.

**Definition 2.4**

A combination $z$ is said to occur in the data set $Y_i$ if the data set contains at least one response corresponding to the combination.

**Definition 2.5**

If $A$ and $B$ are two factors of a structure, and for every level $b$ of
of B there is a level a of A such that \( b \leq a \), then factor A is said to nest factor B, or factor B is said to be nested in A.

Definition 2.6

Factor A is said to directly nest factor B if:

(i) A nests B, and

(ii) there does not exist factor C such that A nests C, and C nests B.

Definition 2.7

A factor A is said to be multiply nested in a set of factors S if and only if for every level a of A there is a combination s of S such that a \( \leq s \).

Definition 2.8

If A and B are two factors of a structure such that A does not nest B and B does not nest A, then the factors A and B are said to be crossed.

Definition 2.9

Let Z be the set of factors in a structure. If for every subset of factors X such that \( X \subseteq Z \), every possible combination \( x \) of X occurs, then the structure is called complete.

Definition 2.10

Let Z be the set of factors in a structure. If every combination that occurs contains the same number of responses, then Z is called a balanced set of factors. If every subset of the set of factors in a structure forms a balanced set, then the structure is called a balanced structure.

Definition 2.11

Let Z be a set of factors in a structure. If Z contains every factor
in which any factor in Z is nested then Z is called an admissible set.

Definition 2.12

If Z is an admissible set of factors and z is an occurring combination, the arithmetic mean of all responses contained in z is an admissible mean.

Definition 2.13

If Z is an admissible set of factors and W is the set of factors in Z each of which nests no other factor in Z other than itself, then W is called the right-most bracket of Z.

Definition 2.14

A graph of a response structure or structure diagram represents each factor of the structure by a small circle and contains a descending line from A to B for every two factors A and B such that A nests B.
An important consideration often overlooked in the design of a general program is ease of problem specification by the user. The minimum sufficient information required by any program for computing the entries in an analysis of variance table, including the expected mean squares, for any balanced complete data set is:

a) the structure of factors,

b) the levels of each factor in the sample and in the population.

The literature contains varied representations for factorial structures, six of which are exemplified in Figure 2.1. These were chosen for comparison because they possess the required level of generality and are each feasible for computer input. They differ, however, in the type of hardware and programming required for their respective interpretation. Form (a), due to Zyskind 1958 (36), is very similar to form (d) due to Nelder 1965 (21). Zyskind used a semi-colon to denote nesting relationships between factors and parenthesis to denote crossed factors, while Nelder used an arrow to denote the nesting operator, a multiplication sign for crossed factors with suitable parenthesis to denote order.

Section J of this chapter comments in detail on a difficulty that arises from the use of this convention, namely, that some structures do not admit this type of representation with a single use of each factor symbol.

The Hasse diagrams used by Throckmorton (Definition 2.14), exemplified by form (b), if used for input to a computer would require special hardware such as a light pen on a cathode ray screen and extensive software for
conversion to digits. The standard model depicted by form (c), is currently used by the AARDVARK (20) program for specification of structures. This form of specifying structure to a computer is unique to AARDVARK and has tremendously important advantages. First, it is natural for statisticians to use the standard model convention, and second, it gives complete flexibility for specification of interaction model terms of interest. The disadvantages of this form of specification are: non-statisticians sometimes find it difficult; it requires extensive programming to scan and interpret the model; it can mislead users because non-specified interactions are automatically pooled with error.

Ultimately, it is necessary that any such representation of structure be converted to the computer's basic unit, namely the binary digit. This is done in all existing programs by varying algorithms and at varying levels of efficiency in programming and core utilization. Thus, computationally the bit matrix forms given by (f) offer the minimum conversion problem for any binary base computer. However, they are not as natural for problem specification by the user as the forms (a) or (e). Form (e) consists of a simple list of factor names with rightmost bracket subscripts used in accordance with Definition 2.10.
Figure 2.1. Equivalent representations of a factorial structure. A three factor data set, one factor nested in one of two crossed factors.
a. Zyskind (36), (37) 
b. Throckmorton (19), (24)

c. Standard model convention

\[ \text{MODEL, } Y = A(I) + B(IJ) + C(K) + \ldots + B(IIK) \]

d. Nelder (21) 
e. Zyskind in Kempthorne et al. (19)

\[
\begin{bmatrix}
0 & 0 & 1 \\
0 & 1 & 1 \\
1 & 0 & 0
\end{bmatrix},
\begin{bmatrix}
1 & 0 & 0 \\
1 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]

f. Hemmerle (16)
D. Structure Matrices

In the subsequent development heavy use will be made of logical operations on binary vectors. The operators used are standard and have the usual definitions and properties found in any logic text. The definitions are summarized for convenience in the following table, Table 2.1.

Table 2.1 Standard Logical Operators

<table>
<thead>
<tr>
<th>Logical Value</th>
<th>(a) l.u.b. &quot;or&quot;</th>
<th>(b) g.l.b. &quot;and&quot;</th>
<th>(c) Negation</th>
<th>(d) exclusive or</th>
<th>(e) subset p \leq q</th>
<th>(f)</th>
</tr>
</thead>
<tbody>
<tr>
<td>p 0 q</td>
<td>p \lor q</td>
<td>p \land q</td>
<td>\neg p</td>
<td>p \oplus q</td>
<td>q \lor \neg p</td>
<td>f</td>
</tr>
<tr>
<td>0 0 0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 1 1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 0 1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1 1 1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

It is proposed that for ease of specifying structure to the computer, the factors be named in accordance with the mode of Figure 2.1(e) with right-most brackets applied, in accordance with Definition 2.13, to the symbols corresponding to factors nesting no other factors in the factor name. For uniqueness, the names can be ordered by the program or the user in ascending order according to the number of symbols outside the rightmost bracket. Note that factors that are nested in factors other than the mean contain
symbols outside and to the left of the right-most bracket, while factors that are nested in no factor other than the mean have a representation consisting of a single symbol in the right-most bracket.

A simple and efficient conversion to a binary form of structure representations is given in the following definition.

**Definition 2.15**

The structure matrix is the matrix whose $i$-th row is formed by equating the $i$-th distinct symbol encountered in the scan of factor names to the binary equivalent of

$$
\sum_{j=1}^{n} 2^{n-j},
$$

where $n$ is the number of factors, with the logical "and" operator applied to all symbols present in the $i$-th factor name.

Thus, for example, consider the structure defined by the factors named $A$, $B$, $A(B)$, $AB(D)$ in the mode of Figure 2.1 (e). Applying the above definition the following formal equivalences result, where $(n)$ denotes the numerical base,

\[
\begin{bmatrix}
A \\
B \\
C \\
D
\end{bmatrix}
= \begin{bmatrix}
2^2 + 2^1 + 2^0 \\
2^3 + 2^1 + 2^0 \\
2^3 + 2^2 + 2^0 \\
2^3 + 2^2 + 2^1
\end{bmatrix}
= \begin{bmatrix}
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
1 & 1 & 1 & 0
\end{bmatrix}
\]

(2.1)

For each factor named the "Boolean product", $(A \land$ operator see
Table 2.1 (b) of the binary representation of all symbols present in the factor name is formed giving the structure matrix

\[
\begin{bmatrix}
  A \\
  B \\
  A(C) \\
  AB(D)
\end{bmatrix} = \begin{bmatrix}
  (0111) \\
  (1011) \\
  (0111) \land (1101) \\
  (0111) \land (1011) \land (1110)
\end{bmatrix} = \begin{bmatrix}
  0111 \\
  1011 \\
  0101 \\
  0010
\end{bmatrix}
\]

It can be seen that in any structure each factor named admits a unique binary representation which is here used as a row of a matrix, and that to each distinct possible structure of factors there corresponds a unique binary matrix representation.

Computationally, the formation of the structure matrix can be accomplished alternatively by the following two steps:

1. Form \( J_n - I_n \), where \( J_n \) is the matrix of order \( n \) each of whose elements is 1, \( I_n \) is the identity matrix of order \( n \), and \( n \) is the number of factors in the structures,

2. For each nested factor form the boolean product of its binary representation with the binary representation of each factor in which it is nested.

We briefly digress from computational considerations to explore the algebraic properties of the resulting matrix.

E. Lattice Properties

Throckmorton (19), (24) observed that with the inclusion of the mean and the error terms the Hasse diagrams of structures have the properties of a lattice when nesting is treated as the partial ordering
relation ($\leq$). Carney (4), (5) proved that the structure resulting from the inclusion of all admissible interactions is also a lattice with respect to the same relation. Thus, if $M$ is the set of model terms and $\leq$ is the nesting relation, the pair $\{M, \leq\}$ satisfies the following two definitions, given for example in Jacobson (17):

**Definition 2.16**

A partially ordered set is a system consisting of a set $M$ and a relation $\leq$ satisfying:

1. $a \leq a', \forall a \in M$,
2. $a \leq b$ and $b \leq a$ iff $a = b, \forall a, b \in M$,
3. $a \leq b$ and $b \leq c$ implies $a \leq c, \forall a, b, c \in M$.

**Definition 2.17**

A lattice structure is a partially ordered set in which any two elements have a least upper bound (l.u.b.) and a greatest lower bound (g.l.b.).

The binary matrix representation of a structure gives implicit logical meanings to the ordering relation $\leq$, the l.u.b. $\lor$, and g.l.b. $\land$ operators of Definition 2.17, applied to terms in a model as elements of the structure. Let the l.u.b. of any two elements $a$, $b$, in a structure be denoted by $a \lor b$, where the operator $\lor$ applies the Boolean sum to the binary representation of $a$ and $b$. For example if $a = (0 \ 1 \ 1 \ 1)$ and $b = (0 \ 0 \ 1 \ 1)$, then $a \lor b = (0 \ 1 \ 1 \ 1)$ is obtained by successive applications of the truth table 2.1 (a). Analogously, let the g.l.b. of any two elements $a$, $b$, in a structure be denoted by $a \land b$, where the operator $\land$ applies the Boolean product to the binary representations. Thus, for the
values of \( a \) and \( b \) above, \((a \land b) = (0\ 0\ 1\ 1)\) is obtained by successive applications of the truth table 2.1 (b).

It is clear that the \( n \) dimensional vector each of whose components is zero corresponds to the error term in all structures and is the lower bound of all the factors in the structure. This follows because any factor in the structure, say \( x \), nests the error term \( \varepsilon \). Thus, \( \varepsilon \leq x \) holds \( \forall x \in M \). Analogously, the \( n \) dimensional vector each of whose components is unity corresponds to the overall mean model term \( \mu \) in all structures, and is an upper bound of all factors in the sense that it nests all factors. Thus, \( x \leq \mu \) holds \( \forall x \in M \). These two elements are called the 0, 1 elements of the lattice structure respectively.

The order relation \( \leq \), which has been used to represent nesting, can be translated to logical operations by the following.

**Lemma 2.1**

In any structure, two factors \( a, b \) are related by the nesting relation \( a \leq b \) if and only if \( b \lor \neg a \) is true, i.e., the resulting vector contains all one's.

If \( a \leq b \), then whenever the \( i \)-th binary component of \( a \), \( a_i \) say, is unity, the corresponding component of \( b \), \( b_i \) must be unity. Consequently, applying Table 2.1 (c), either \( \neg a_i \) is unity or \( b_i \) must necessarily be unity if \( a \leq b \).

It follows directly and in accordance with standard usage of \( \leq \), \( \land \), and \( \lor \), that if \( a \leq b \) then \( b = a \lor b \) and \( a = a \land b \). That is, if \( a \) is nested in \( b \), then \( a \) and \( b \) are respectively the g.l.b. and l.u.b. of the two elements \( a \) and \( b \).
As a further illustration, consider the factorial structure diagrammatically represented by

\[
\begin{array}{c}
A \\
\downarrow \\
A(0) \\
\downarrow \\
e \\
\uparrow \\
AB(D) \\
\uparrow \\
B \\
\uparrow \\
A \quad C
\end{array}
\]

with resulting structure matrix as given.

Then \( C \preceq A \iff A \lor C \) is true,

since \( A \lor C \equiv (0 1 1 1) \lor (0 1 0 1) \equiv (0 1 1 1) \lor (1 0 1 0) \equiv (1 1 1 1) \);

while \( C \preceq B \not\iff B \lor C \) is true,

since \( B \lor C \equiv (1 0 1 1) \lor (0 1 0 1) \equiv (1 0 1 1) \lor (1 0 1 0) \equiv (1 0 1 1) \),

and the presence of a zero anywhere in the resulting expansion results in a "false" condition.

The following lattice properties of experimental structures are given for reasons of general interest apart from computational significance.

Because the structure of model terms is a lattice the operators must satisfy
(i) \( a \lor b \equiv b \lor a \)
(ii) \( (a \lor b) \lor c \equiv a \lor (b \lor c) \)
(iii) \( a \lor a \equiv a \)
(iv) \( (a \lor b) \land a \equiv a \);

and the duality relationships obtained by the interchange of the two operators. Additionally, a lattice satisfying

(v) \( a \geq b \implies a \land (b \lor c) \equiv b \lor (a \land c) \)

is called a modular lattice.

The modularity of a lattice is known to be a weaker condition than the distributive property of a lattice

(vi) \( a \lor (b \land c) \equiv (a \lor b) \land (a \lor c) \),

which holds trivially for the lattice of model terms of a structure. Conditions (i) to (iv) are consistent with the definitions of Boolean sum and products. The Truth Table method of proof for (v) and (vi) are used in the following.

Theorem 2.1

A factorial structure with the ordering relation \( \leq \), the l.u.b. \( V \), and g.l.b. \( \land \), logically defined on the rows of the resulting binary matrix representation is a distributive lattice.

Proof: We consider the Truth Table for (vi) applied to all the possible assignments of values for the i-th element in the binary representation of three model terms denoted \( a, b, c \):
Since the Truth Table must hold $\forall b_i (\forall i)$ holds for all structures and therefore every structure is a distributive lattice.

**Corollary 2.1**

A factorial structure with the ordering relation $\leq$, the l.u.b. $\lor$, and g.l.b. $\land$, logically defined on the rows of the resulting binary matrix representation is a modular lattice.

<table>
<thead>
<tr>
<th>$a_i b_i c_i$</th>
<th>$a_i \lor (b_i \land c_i)$</th>
<th>$(a_i \lor b_i) \land (a_i \lor c_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 0 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 0 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 1 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0 1 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 0 0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 0 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 1 0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1 1 1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$a_i b_i c_i$</th>
<th>$a_i \land (b_i \lor c_i)$</th>
<th>$b_i \lor (a_i \land c_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) 0 0 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2) 0 0 1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3) 0 1 0</td>
<td>0</td>
<td>$\neq$ 1</td>
</tr>
<tr>
<td>4) 0 1 1</td>
<td>0</td>
<td>$\neq$ 1</td>
</tr>
<tr>
<td>5) 1 0 0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6) 1 0 1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7) 1 1 0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>8) 1 1 1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Thus, for the Boolean operators defined, the consequence of the condition (v) is violated only in two of the eight possible combinations of logical assignment of values to $a_i$, $b_i$, $c_i$. However, in accordance with the definitions for representations of factors, the requirement $a \geq b$ of condition (v) implies that if $b_i = 1$, necessarily $a_i = 1$. Therefore, since 3 and 4 violate the antecedent of the implication the corollary is proved.

A useful property of modular lattices is the chain condition of Schreier's theorem in (17), namely that any two finite descending chains connecting two elements of a modular lattice have equivalent refinements. Additionally, we have the following theorem.

Theorem 2.2

For any factorial structure, if factor $a$ nests factor $b$, and for any other factor $c$, $a \lor c = b \lor c$, $a \land c = b \land c$, then $a \equiv b$.

Proof: Let $a = (a_1, a_2, ..., a_n)$, $b = (b_1, b_2, ..., b_n)$, and $c = (c_1, c_2, ..., c_n)$, where the $a_i$'s and $b_j$'s are the binary digits in the structure matrix representation. Since $a$ nests $b$ we have $a_i = 1$ whenever $b_i = 1$. Therefore, we consider the truth tables for the six remaining cases.

<table>
<thead>
<tr>
<th>$a_i$</th>
<th>$b_i$</th>
<th>$c_i$</th>
<th>$a_i \lor c_i$</th>
<th>$b_i \lor c_i$</th>
<th>$a_i \land c_i$</th>
<th>$b_i \land c_i$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>$\neq$</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>$\neq$</td>
</tr>
<tr>
<td>5)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
The conditions of the theorem are not satisfied by 3, 4, while 1, 2, 5, 6 satisfy $a_i = b_i$. Since the truth table must hold for all $i$ the result is proved.

These properties of modular lattices justify an obvious algorithm for forming all admissible interactions from any structure matrix, namely, that the boolean product of the rows yield all admissible model terms. If, in the previous example, the Boolean product (equivalent in our content the greatest lower bound) of the rows of the structure matrix (binary factorial representation) are taken first, two at a time, then three at a time, etc., to $n$ at a time, and if repetitions are disregarded we obtain from the structure matrix

\[
\begin{bmatrix}
\mu \\
A \\
B \\
A(C) \\
AB(D) \\
E
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 \\
0 & 1 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

(2.3)

the matrix

\[
\begin{bmatrix}
\mu \\
A \\
B \\
AB \\
A(C) \\
A(B) \\
AB(D) \\
AB(\text{CD}) \\
E
\end{bmatrix}
\begin{bmatrix}
\mu \\
A \\
B \\
A \land B \\
A(C) \\
B \land A(C) \\
AB(D) \\
A(C) \land AB(D) \\
E
\end{bmatrix}
\begin{bmatrix}
1 & 1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 \\
0 & 0 & 1 & 1 & 1 \\
0 & 0 & 0 & 1 & 1 \\
0 & 0 & 1 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

(2.4)
For convenience, term this the full structure matrix. Its properties are:

(i) each row corresponds uniquely to an admissible set (see Definition 2.11),

(ii) the decimal conversion of each row provides a convenient unique numerical tag for storage and retrieval of all computations pertaining to that model term,

(iii) the model terms for computer annotation of the analysis of variance table are obtained directly by symbolic concatenation of names for factors whose Boolean product contributed to the formation of that model term.

Note that the full structure matrix (2.4) is completely implied by the structure matrix (2.2), which in turn is completely implied by the adoption of a naming convention for the factors.

Whenever the boolean product operation on the matrix in (2.3) yields an admissible interaction, the same operation is applied to the corresponding rows of the matrix $J_n - I_n$. The resulting matrix

\[
\begin{bmatrix}
\mu \\
A \\
B \\
A \land B \\
C \\
B \land C \\
D \\
CD \\
E
\end{bmatrix} = \\
\begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 \\
1 & 1 & 0 & 1 \\
1 & 0 & 0 & 1 \\
1 & 1 & 1 & 0 \\
1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]
contains 0's in positions corresponding to the "rightmost bracket" as defined by Zyskind (see Definition 2.13). The computation of components of variation for balanced complete structures will be obviously facilitated by the availability of this matrix which will be subsequently termed the rightmost bracket matrix.

F. Admissible Means

The means are an essential preliminary reduction of a structured data set. Efficient condensation of data sets to the admissible partial means (see Definition 2.12) for arbitrary structures of arbitrarily large size, is a computational problem meriting careful consideration. Because the means are vital in extracting information from data beyond that information given in A.O.V. tables, it is essential that the algorithm

(i) be efficient computationally for large or small data sets,

(ii) minimize required core storage,

(iii) provide a natural mode of identifying means for subsequent retrieval and use by other algorithms.

The first requirement cannot seemingly be satisfied by a single procedure. A data set of 250,000 elements must necessarily be processed differently than one of, say, 100 elements. The computation of means and components in core, although extremely fast, requires approximately 5 times N words of core storage, where N is the number of observations in a data set. Because this is frequently excessive, existing algorithms which do not differentiate on size are limited on size or are inefficient in
computations. To minimize core utilization, the algorithm to be described relies on a direct accessing scheme and the availability of peripheral storage devices. Convenient and conventional annotation for output of partial means can be obtained from the rows of the full structure matrix, where the ones are interpreted as indices of summation. The uniqueness of the rows provide the key for direct access of sets of means. For instance, in the preceding example the successive rows of the matrix given in (2.4) correspond respectively to \( y_{1}, y_{2}, \ldots, y_{j-1}, y_{j}, \ldots \), \( y_{i,k}, y_{j,k}, \ldots, y_{i,j}, y_{ijkl}, y_{ijklm} \).

The objective is to form all the required admissible partial means one model term at a time, utilizing in core only the available condensations of the data represented by previously computed means for other model terms. The algorithm must provide a method of identifying and locating in core every value entering each mean. The essential control information is contained in the following four vectors which, in their respective \( i \)-th positions, contain the following for the computation of the \( i \)-th string of means.

- \( U(i) \) = the identification of the string of means to be directly accessed from which the means for the \( i \)-th model term will be computed.
- \( S(i) \) = the range of the subscript of summation.
- \( K(i) \) = the distance between two elements entering each mean.
- \( J(i) \) = the number of partial means for the \( i \)-th model term.

**Lemma 2.2**

A necessary and sufficient condition that the means for a model term
be constructible from the means for model term $a$ is: $(a < b)\\n(= (b \lor \neg a))$.

That is, $b$ must be the l.u.b. of $a$ and $b$. Clearly, all admissible means, for example, could be computed from the set of all observations which in this context correspond to the error term $\varepsilon$, while conversely the grand mean $\mu$ could be computed from any other set of admissible partial means. This follows clearly because $\varepsilon \leq x \lor x \in M$, and, by the previous definitions, $x \lor \neg \varepsilon \equiv \neg \varepsilon \equiv \mu$ which has a representation of all one's, while $x \leq \mu$, $\forall x \in M$ and $\mu \lor \neg x \equiv \mu$.

PL-1 allows simple definition of complex logical connectives by a convenient convention. Thus, the truth table for $b \lor \neg a$ is given in Table 2.1(e) and the appropriate PL-1 statement would be

$$\text{IF} (\text{BOOL}(A, B, '11101')) \text{ THEN } \ldots \ldots .$$

When the number of factors, model terms and observations are large, significant savings are achieved by summing on one subscript only. This consideration is described by the following Lemma.

**Lemma 2.3**

Let $a, b$ satisfy Lemma 2.2, i.e. $a \leq b$. If $\text{SUM}(c)$ denotes the sum of digits in any binary number $c$, then $\text{SUM}(b \land \neg a) = 1$ implies that the means for $b$ are obtainable from those of $a$ by summation on a single subscripts.

The condition $\text{SUM}(b \land \neg a) = 1$ can be equivalently expressed by saying $b$ is a "cover" of $a$, in the sense that $b \geq a$ and there exists no $c$ such that $b \geq c \geq a$. Alternatively, we can say that $a$ is a direct nestor of $b$ in the full structure matrix (see Definition 2.6). The chain conditions
of a finite partially ordered set guarantees that for any model term, \( a \), Lemma 2.3 will be satisfied by some model, \( b \). Note that the appropriate truth table is given in table 1, and that the function \( \text{SUM} \) is standard in the PL-1 language with precisely the meaning defined so that the above procedure is easily programmed by a PL-1 instruction containing:

\[
\text{IF}(\text{SUM}(\text{BOOL}(A, B, '0 1 0 0')) = 1) \text{ THEN } \ldots \ldots
\]

The algorithm requires examination of the full structure matrix and storage in a vector, here referred to as \( U \), for each model term, the identification of the model term satisfying Lemmas 2.2 and 2.3. For illustration, considering the previous example, the entries in \( U \) would be as in Figure 2.2.

Let \( R \) be an array containing the number levels observed on each factor ordered such that \( R(I) \) is the range of the subscript corresponding to the factor of the I-th column of the full structure matrix.

**Lemma 2.4**

The inner product of the vector \( c = b \Lambda^T a \) satisfying Lemma 2.3, with \( R \), the vector of subscript ranges, gives \( S \), the vector containing the range of the subscript of summation, i.e., \( S = c'R \). For example, if \( R = (2, 3, 5, 7, 4) \) the resulting vector \( S \) would be as shown in Figure 2.2.

The balanced complete property of the data and the convention of storing observations lexicographically with respect to indices allows convenient location of elements entering each partial mean. Because the elements entering the summations in the computation of a string of means do not necessarily occur in adjacent positions, the program requires a vector of integers, \( K \), which provides the distance between successive elements.
Let $M$ be a binary square matrix whose $i$-th row is the binary equivalent of

$$M(I) = \sum_{j=1}^{n} 2^{n-j},$$

where $n$ is the order of the structure. For example if $n = 5$ then

$$M = \begin{bmatrix}
2^4 & & & & \\
2^4 + 2^3 & & & \\
2^4 + 2^3 + 2^2 & & \\
2^4 + 2^3 + 2^2 + 2^1 & & \\
2^4 + 2^3 + 2^2 + 2^1 + 2^0 & & & & \\
\end{bmatrix} = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 \\
& 1 & 1 & 0 & 0 \\
& & 1 & 1 & 1 \\
& & & 1 & 1 \\
& & & & 1 & 1 \\
\end{bmatrix}. $$

For each vector $c$ in Lemma 2.4, the PL-1 function $N = \text{INDEX}(c,'1')$ returns to $N$, a value corresponding to the position from the left of the 1-bit in $c$. For example, if $c = (01000)$, then $N = 2$.

**Lemma 2.5**

The distance between two elements entering a summation in a partial mean of model term $b$ is given by

$$K = \prod \left[ \left( \Pi(N) \lor b \right) \Theta R \right],$$

where $\Theta$ denotes the componentwise product and $\Pi$ the product of non-zero elements of a vector.
<table>
<thead>
<tr>
<th>Model Term</th>
<th>Binary Representation</th>
<th>Decimal Conversion</th>
<th>Binary Representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>b</td>
<td>a</td>
<td>c = b∧¬a</td>
</tr>
<tr>
<td>1) μ</td>
<td>11111</td>
<td>31</td>
<td>01111</td>
</tr>
<tr>
<td>2) A</td>
<td>01111</td>
<td>15</td>
<td>00111</td>
</tr>
<tr>
<td>3) B</td>
<td>10111</td>
<td>27</td>
<td>00111</td>
</tr>
<tr>
<td>4) AB</td>
<td>00111</td>
<td>7</td>
<td>00011</td>
</tr>
<tr>
<td>5) A(C)</td>
<td>01011</td>
<td>11</td>
<td>00011</td>
</tr>
<tr>
<td>6) A(BC)</td>
<td>00011</td>
<td>3</td>
<td>00001</td>
</tr>
<tr>
<td>7) AB(D)</td>
<td>00101</td>
<td>5</td>
<td>00001</td>
</tr>
<tr>
<td>8) AB(CD)</td>
<td>00001</td>
<td>1</td>
<td>00000</td>
</tr>
<tr>
<td>9) E</td>
<td>00000</td>
<td>0</td>
<td>00000</td>
</tr>
</tbody>
</table>

Figure 2.2. Resulting entries in the U and S vectors controlling computation of admissible partial means.
Figure 2.3. Logic of subroutine for construction of all admissible means
Construction of Means for Model Term

is Means (U(i)) in CORE?

Yes

No

Read MEANS (U(i))

\[ \text{Mean } (j) = \text{Mean } (j) + \text{Core } (j+k(i) \cdot (i) \text{)} \]

\[ \text{Mean } (j) = \text{Mean } (j) + \text{Core } (j+k(i) \cdot (k-l)) \]

j = 1
k = 1

is \( j \geq k(I) \)

Yes

Write Means(I)

Increment i

No

j = j + 1

is \( k \geq s(i) \)

Yes

No

k = k + 1
Lemma 2.6

The number of partial means corresponding to a given model term \( b \) is

\[ J = (\mathbf{1} b)'R. \]

G. Components of Variation

Let \( \mathcal{J} = \{s_1, s_2, s_3, \ldots, s_m\} \) denote the set of rows of the full structure matrix, and \( \mathcal{R} = \{r_1, r_2, r_3, \ldots, r_k\} \) the corresponding rows of the rightmost bracket matrix. The set of admissible means in the experimental structure is given formally by the set of \( y_{s_i} \) where the presence of 1's in the expansion \( s_i \) denote indices of summation and the 0's denote variable subscripts over sample ranges. To each admissible mean there corresponds linear combinations of admissible means known as components. The components have various important properties which were given by Zyskind (36). Computationally, the components define the relation used in decomposing the total sum of squares. As such, the sum of squares of individual components over the range of indices are the usual sums of squares in A.O.V. tables.

The admissible means entering the linear combination with non-zero coefficients have been defined by Zyskind (see Definition 2.12), as those means which can be obtained from the term in question by omission of subscripts only from the rightmost bracket. Whenever an odd number of indices is omitted the mean is preceded by a negative sign, and whenever an even number is omitted the mean is preceded by a positive sign, the number zero here considered even. These rules define a relation,
\[ \beta = \gamma y, \]

giving the vector of components \( \beta \) in terms of the corresponding vector of admissible means \( y \).

Programming the construction of the square matrix \( \gamma \) is made especially simple by the logical development thus far obtained. The matrix \( \gamma \) is known to have the following special properties,

(i) each element is either 0, 1 or -1,

(ii) the sum of each row, except that corresponding to the mean, is zero,

(iii) the diagonal elements are equal to one,

(iv) the matrix is lower triangular.

The elements \( \gamma_{ij} \) in terms of programmable logical operations are given by:

\[ \gamma_{ij} = (-1)^{S(\langle \neg (s_j \leq s_i) \rangle \neg (s_i \leq s_j) \langle \neg (s_j \leq s_i) \rangle)} , \quad i \neq j \]

where \( s_k \in \beta \), \( r_k \in \beta \) denote the \( k \)-th rows of the full structure matrix and rightmost bracket matrices respectively.

The operator \( S(\langle \rangle) \), which as previously defined is \( S(\langle \rangle) = \delta^\dagger \delta \), simply counts the bits in a binary string. The expression (2.7) has the following explanation. The "omission of an index" requirement is satisfied iff \( s_i \leq s_j \), or equivalently, if the model term \( s_i \) is a nestor of \( s_j \) which implies \( s_j \lor \neg s_i \) has truth value '1'. When this condition is logically true, the indices omitted are given by the zeroes in binary expansion of the reverse inequality \( s_j \leq s_i \), which may be obtained by evaluating \( s_i \lor \neg s_j \). The requirement that these indices omitted be "rightmost
bracket 'only' is satisfied iff \( r_1 \leq (s_j \leq s_1) \), which implies \((s_j \leq s_i) \lor \neg r_1 \) has the truth value '1', thus giving expression (2.7) a non-zero value. It is clear that the "odd, even" condition refers to the number of 0's in the expansion of \( s_j \leq s_i \); hence, the appropriate sign with which a non-zero coefficient enters is given by summing the bits of the complement resulting in the term \((-1)^{S(s_j \leq s_i)}\).

By applying the sentential calculus and the definition of the nesting relation \( \leq \), expression (2.7) may be rewritten:

\[
(2.8) \quad x_{ij} = (-1)^{S((s_i \lor \neg s_j) \land (s_i \lor \neg s_i) \land \neg r_i)} \quad i > j.
\]

Thus the expression (2.8) requires only the \( \land \), \( \lor \), and \( \neg \) operators which are logical conjunctions in almost any high-level programming language such as COBOL, FORTRAN, and PL/1.

The elements of \( y^{-1} \) giving the means as linear combinations of the components by the matrix equation

\[
y = y^{-1} \beta,
\]

are obtained simply as:

\[
x_{ij} = \begin{cases} 
1 & \text{if } s_i \leq s_j, \forall i \geq j \\
1 & \text{if } i = j \\
0 & \text{otherwise}.
\end{cases}
\]

Computationally, it is almost as convenient to obtain \( y \) from \( y^{-1} \) by a back substitution, but either procedure can be programmed efficiently in terms of computer time and core utilization.

Figure 2.4 illustrates the operations defined by (2.7) for the
structure of the example. The appropriate structure matrices may be found in Section 2, (2.4), and (2.5). By conditions (iii) and (iv) above, the elements of $\gamma_{ij}$ are required only for $i > j$ as given in Figure 2.4.

The resulting linear combinations of means defining the components are:

\[
\begin{align*}
\beta_{11111} &= y(11111) \\
\beta_{01111} &= -y(11111) + y(01111) \\
\beta_{10111} &= -y(11111) + y(10111) \\
\beta_{00111} &= y(11111) - y(01111) - y(10111) + y(00111) \\
\beta_{01011} &= -y(01111) + y(01011) \\
\beta_{00011} &= y(01111) - y(00111) - y(01011) + y(00111) \\
\beta_{00001} &= y(00011) - y(00101) + y(00001) \\
\beta_{00000} &= y(00001) - y(00000) \\
\end{align*}
\]

An important relation exemplified by the components above is the identity decomposition of a typical observation into the sum of components in the structure:

\[
y(00000) = \sum_{x \in B} \beta_x x.
\]

As a consequence, the sum of squares in the A.O.V. table corresponding to any given line, uniquely identified by $x \in B$, is given by

\[
SS_x = \sum_{1 \leq x_i \leq r_i} \beta^2 (x_1, x_2, \ldots, x_n),
\]

where $r_i$ is range of the $i$-th subscript as contained in the vector $A$. That
is, the \( SS_x \) is the sum of squares of corresponding components over the range of all subscripts.

**H. Expected Mean Squares**

To each \( s_i \in \mathcal{S} \) there corresponds a linear function of the components of variation \( \sigma^2_x, x \in \mathcal{S} \). These linear functions were called \textit{cap sigmas}, \( \Sigma \)'s, and were introduced by Wilk (27), (28) to simplify the expressions for expected mean squares in certain analysis of variance tables. A general definition of the \( \Sigma \)'s, given by Zyskind (36) admits a formulation in terms of logical operations on the elements of \( \mathcal{S} \) and \( \mathcal{R} \), and results in a simple computational algorithm for the expected mean squares for the balanced complete structures here considered.

Let \( R_p \) denote the vector containing the levels of each factor in the population, ordered the same as the vector \( R \) of sample subscript ranges previously defined. The components of variation \( \sigma^2_x \) entering the expression for \( \Sigma_{s_i} \) are those that satisfy

\[
(2.9) \quad (x \leq s_i) \land [r_x \leq (s_i \leq x)].
\]

In the terminology of Zyskind's general definition, the condition that "the subscripts of \( \sigma^2 \) include the set of subscripts corresponding to the leading term of the component as a subset" is satisfied iff \( x \leq s_i \).

Further, the condition that "excess subscripts", here given by the 0's in the expression \( s_i \leq x \), belong exclusively to the rightmost bracket" is satisfied iff \( r_x \leq (s_i \leq x) \).
<table>
<thead>
<tr>
<th>i,j</th>
<th>( s_i \leq s_j )</th>
<th>( s_j \leq s_i )</th>
<th>( r_i \leq (s_j \leq s_i) )</th>
<th>( S( (s_j &lt; s_i) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>21</td>
<td>11111</td>
<td>01111</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>31</td>
<td>11111</td>
<td>10111</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>32</td>
<td>01111</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>11111</td>
<td>00111</td>
<td>11111</td>
<td>2</td>
</tr>
<tr>
<td>42</td>
<td>11111</td>
<td>10111</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>43</td>
<td>11111</td>
<td>01111</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>51</td>
<td>11111</td>
<td>01011</td>
<td>01111</td>
<td>-</td>
</tr>
<tr>
<td>52</td>
<td>11111</td>
<td>11011</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>53</td>
<td>10111</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>10111</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61</td>
<td>11111</td>
<td>00011</td>
<td>01111</td>
<td>-</td>
</tr>
<tr>
<td>62</td>
<td>11111</td>
<td>10011</td>
<td>11111</td>
<td>2</td>
</tr>
<tr>
<td>63</td>
<td>11111</td>
<td>01011</td>
<td>01111</td>
<td>-</td>
</tr>
<tr>
<td>64</td>
<td>11111</td>
<td>11011</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>65</td>
<td>11111</td>
<td>10111</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>71</td>
<td>11111</td>
<td>00101</td>
<td>00111</td>
<td>-</td>
</tr>
<tr>
<td>72</td>
<td>11111</td>
<td>10101</td>
<td>10111</td>
<td>-</td>
</tr>
<tr>
<td>73</td>
<td>11111</td>
<td>01101</td>
<td>01111</td>
<td>-</td>
</tr>
<tr>
<td>74</td>
<td>11111</td>
<td>11101</td>
<td>11111</td>
<td>1</td>
</tr>
<tr>
<td>75</td>
<td>11011</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76</td>
<td>11011</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.4. Construction of \( \gamma \) for the four factor model with replication of the example
<table>
<thead>
<tr>
<th>i</th>
<th>( s_i \leq s_j )</th>
<th>( s_j \leq s_i )</th>
<th>( r_i \leq (s_j \leq s_i) )</th>
<th>( S(s_i \lor s_j) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1 1 1 1 1</td>
<td>0 0 0 0 1</td>
<td>0 0 1 1 1</td>
<td>——</td>
</tr>
<tr>
<td>2</td>
<td>1 1 1 1 1</td>
<td>1 0 0 0 1</td>
<td>1 0 1 1 1</td>
<td>——</td>
</tr>
<tr>
<td>3</td>
<td>1 1 1 1 1</td>
<td>0 1 0 0 1</td>
<td>0 1 1 1 1</td>
<td>——</td>
</tr>
<tr>
<td>4</td>
<td>1 1 1 1 1</td>
<td>1 0 0 0 1</td>
<td>1 1 1 1 1</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1 1 1 1 1</td>
<td>0 1 0 0 1</td>
<td>1 0 1 1 1</td>
<td>——</td>
</tr>
<tr>
<td>6</td>
<td>1 1 1 1 1</td>
<td>1 1 0 0 1</td>
<td>1 1 1 1 1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1 1 1 1 1</td>
<td>1 1 0 1 1</td>
<td>1 1 1 1 1</td>
<td>1</td>
</tr>
<tr>
<td>8</td>
<td>1 1 1 1 1</td>
<td>0 0 0 0 0</td>
<td>0 0 0 0 1</td>
<td>——</td>
</tr>
<tr>
<td>9</td>
<td>1 1 1 1 1</td>
<td>1 0 0 0 0</td>
<td>1 0 0 0 1</td>
<td>——</td>
</tr>
<tr>
<td>10</td>
<td>1 1 1 1 1</td>
<td>0 1 0 0 0</td>
<td>0 1 0 0 1</td>
<td>——</td>
</tr>
<tr>
<td>11</td>
<td>1 1 1 1 1</td>
<td>1 1 0 0 0</td>
<td>1 1 0 0 1</td>
<td>——</td>
</tr>
<tr>
<td>12</td>
<td>1 1 1 1 1</td>
<td>1 0 1 0 0</td>
<td>1 0 1 0 1</td>
<td>——</td>
</tr>
<tr>
<td>13</td>
<td>1 1 1 1 1</td>
<td>1 1 0 0 0</td>
<td>1 1 0 0 1</td>
<td>——</td>
</tr>
<tr>
<td>14</td>
<td>1 1 1 1 1</td>
<td>1 1 0 1 0</td>
<td>1 1 0 1 1</td>
<td>——</td>
</tr>
<tr>
<td>15</td>
<td>1 1 1 1 1</td>
<td>1 1 1 0 0</td>
<td>1 1 1 0 1</td>
<td>——</td>
</tr>
<tr>
<td>16</td>
<td>1 1 1 1 1</td>
<td>1 1 1 1 0</td>
<td>1 1 1 1 1</td>
<td>1</td>
</tr>
</tbody>
</table>

\[ \gamma = \begin{bmatrix} 1 \\ -1 & 1 \\ -1 & 0 & 1 \\ 1 & -1 & -1 & 1 \\ 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & -1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \quad \gamma^{-1} = \begin{bmatrix} 1 \\ 1 \\ 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \]
The appropriate coefficient for each \( \sigma^2_x \) entering the definition of \( \Sigma_{s_i} \) was given by

\[
(\neg)^k
\prod \text{Product of population ranges of the excess indices}
\]

where \( k \) is the number of excess subscripts. This coefficient may be computed as

\[
\frac{(\neg)^S(\gamma(s_i \leq x))}{\prod (\gamma(s_i \leq s_i) \in R_p)}
\]

where \( \gamma \) denotes the usual \( \gamma \)-l component-wise product of two vectors and \( \gamma \) denotes the function which takes the product of all non-zero elements of a vector. The logical definition of a \( \Sigma \) can then be given as

\[
(2.10) \quad \Sigma_{s_i} = \sum_{s_i \in S} \frac{(\neg)^S(\gamma(s_i \leq x)) \sigma^2_x}{\prod (\gamma(s_i \leq s_i) \in R_p)} [(x \leq s_i) \land (x \leq (s_i \leq x)].
\]

Again, operating on the expression in (2.10), one equivalently obtains

\[
(2.11) \quad \Sigma_{s_i} = \sum_{s_i \in S} \frac{(\neg)^S(\gamma(x \land s_i)) \sigma^2_x}{\prod (\gamma(x \land s_i) \in R_p)} [(s_i \lor x) \land [(x \lor s_i) \lor \neg x].
\]

That the leading term of the component \( s_i \) is \( \sigma^2_{s_i} \) can be seen by letting \( x = s_i \) in (2.11); for then the conditional expression becomes

\[
(s_i \lor \neg s_i) \land (s_i \lor \neg s_i \lor \neg x),
\]

which, obviously, is always true for any expansion of \( s_i \). For an illustration of the computational results of applying (2.11), see Figure 2.5.

For balanced complete structured data from balanced complete identically structured conceptual populations, Throckmorton has shown that
sample analogues of the population \( \Sigma' \)'s satisfy the important property of inheritance on the average

\[
E(\Sigma'_{s_i}) = \Sigma_{s_i},
\]

where the ' is used to denote the fact that the \( \Sigma'_{s_i} \) are sample random variables.

When the conduct of the experiment involves the physical act of randomization, the above one-to-one correspondence of population and sample structures need not hold, as, for example, in the general randomized block experiments. All existing general purpose analysis of variance programs known to the author do not directly treat such cases, leaving to the interpreter of the outputs the decision on appropriate tests based on their own computation of expected mean squares. However, by inputs of the two structures simple expressions for the expectations of mean squares can be obtained in terms of \( \Sigma^E \) quantities defined by Throckmorton.

Let \( \mathcal{F} \) denote the set of rows of the population structure matrix.

Then, \( s_i \in \mathcal{F} \)

\[
\Sigma^E_{s_i} = \sum_{x \in \mathcal{F}} \Sigma x [ (\forall x \leq r_i) \land (s_i \leq s)].
\]

The condition in (2.13) can equivalently be written,

\[
(r_i \lor \exists x) \land (x \lor \exists s_i).
\]
Figure 2.5. Illustration of logical operators for construction of $\Sigma^{(00111)}$ in terms of components of variation for the structure of the example.
\[ \Sigma(00111) = \sum_{x \in \mathbb{R}_{p}} (-1)^{\left((11000) \land x\right)} \sigma^2_{x} \times \left([x \lor (11000) \lor \neg r_{x}] x \lor (11000) \lor \neg r_{x}\right) \]

<table>
<thead>
<tr>
<th>(x)</th>
<th>((00111) \lor \neg x)</th>
<th>(r_{x})</th>
<th>(x \lor (11000) \lor \neg r_{x})</th>
<th>(r_{x} \land (00111))</th>
<th>(S(r_{x} \land (00111)))</th>
</tr>
</thead>
<tbody>
<tr>
<td>11111</td>
<td>00111</td>
<td>Stop</td>
<td>11111</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>01111</td>
<td>10111</td>
<td>Stop</td>
<td>01111</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>10111</td>
<td>01111</td>
<td>Stop</td>
<td>10111</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>00111</td>
<td>11111</td>
<td>00111</td>
<td>11111</td>
<td>00000</td>
<td>0</td>
</tr>
<tr>
<td>01011</td>
<td>10111</td>
<td>Stop</td>
<td>11011</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td>00011</td>
<td>11111</td>
<td>10011</td>
<td>11111</td>
<td>00100</td>
<td>1</td>
</tr>
<tr>
<td>00101</td>
<td>11111</td>
<td>11101</td>
<td>11111</td>
<td>00010</td>
<td>1</td>
</tr>
<tr>
<td>00001</td>
<td>11111</td>
<td>11001</td>
<td>11111</td>
<td>00110</td>
<td>2</td>
</tr>
<tr>
<td>00000</td>
<td>11111</td>
<td>11110</td>
<td>11001</td>
<td>Stop</td>
<td>-----</td>
</tr>
</tbody>
</table>

\[ \cdot \cdot \cdot \Sigma(00111) = \sigma^2_{\Sigma(00111)} - \frac{\sigma^2_{(00011)}}{R_{p}(3)} - \frac{\sigma^2_{(00101)}}{R_{p}(4)} - \frac{\sigma^2_{(00001)}}{R_{p}(3) \cdot R_{p}(4)} \]

Corresponding
Conventional
National
\[ \Sigma_{AB} = \sigma^2_{AB} - \frac{\sigma^2_{A(BC)}}{C} - \frac{\sigma^2_{AB(D)}}{D} - \frac{\sigma^2_{AB(CD)}}{CD} \]
For example, the general randomized block experiment has population structure

\[
\begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

and observational structure

\[
\begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
0 & 0 & 1 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 \\
\end{bmatrix}
\]

and observational structure
Consider $\Sigma^* = \Sigma_{s_i}^{(0001)}$, and let $p_i$ denote the $i$th row of $P$. The following computations are performed:

\[ r_1 \lor p_1 = (1101) \lor (0000) = (1101) \quad \therefore \Sigma \text{ does not enter}, \]

\[ r_1 \lor p_2 = (1101) \lor (1000) = (1101) \quad \therefore \Sigma \text{ does not enter}, \]

\[ r_1 \lor p_3 = (1101) \lor (0100) = (1101) \quad \therefore \Sigma \text{ does not enter}, \]

\[ r_1 \lor p_4 = (1101) \lor (1100) = (1101) \quad \therefore \Sigma \text{ does not enter}, \]

\[ r_1 \lor p_5 = (1101) \lor (1010) = (1111), \]

and

\[ p_5 \lor s_1 = (0101) \lor (1110) = (1111) \quad \therefore \Sigma \text{ enters}, \]

\[ r_1 \lor p_6 = (1101) \lor (1110) = (1111) \]

and

\[ p_6 \lor s_1 = (0001) \lor (1110) = (1111) \quad \therefore \Sigma \text{ enters}, \]

\[ r_1 \lor p_7 = (1101) \lor (1111) = (1111) \]

and

\[ p_7 \lor s_1 = (0000) \lor (1110) = (1110) \quad \therefore \Sigma \text{ does not enter}. \]

Hence, in accord with known results,

\[ \Sigma^*_{s_i}^{(0001)} = \Sigma_{s_i}^{(0101)} + \Sigma_{s_i}^{(0001)} . \]

In general, for balanced populations, Throckmorton developed a scheme of combining the two structures diagrammatically into

\[ \text{Diagram} \]

\[ B \xrightarrow{\Sigma} T \]

\[ p \]

\[ e \]
The expected mean squares of any given model can then be expressed in terms of sample $\Sigma^*$'s by

$$EMS_{s_i} = \sum_{x \in J} \pi[x \cap R] \Sigma^* [x \leq s_i].$$

In those cases where the sample and population structures are identical, there remain computational considerations for simplification of the formulas for E.M.S. These simplifications arise when one or more elements of $R_p$ are infinite, corresponding to a sampling of factor levels from an infinite population as well as when one or more elements in $R_p$ equal corresponding elements of $R$, which is the case usually referred to "fixed" effects.

The vectors $R$ and $R_p$ of sample and population ranges, respectively, have been assumed available inputs. From a programming viewpoint, it is conceivable to require $R_p(I) = 0$ if the population range of its $i$-th subscript is infinite. Clearly, a value 0 in this vector has no meaning other than that which one may assign. The algorithmic procedure would be to construct the following vectors:

$$\text{fixed effects} = f(I) = \begin{cases} 1 & \text{if } R_p(I) = 0 \\ 0 & \text{otherwise} \end{cases},$$

$$\text{random effects} = d(I) = \begin{cases} 1 & \text{if } R_p(I) = 0 \\ 0 & \text{otherwise} \end{cases},$$
semi-random = m(I) = (1,1, ...,1) - f - d ,

correction factors = c(I) = \begin{cases} 
1 - \frac{R(I)}{R_p(I)} & \text{if } m(I) = 1 \\
\bar{d}(I) & \text{otherwise}
\end{cases}

Let $E_{s_i}$, $s_i \in S$, denote the expected mean square in question. The following known results are converted to boolean operations providing all the necessary corrections for "fixed" factors and factors sampled from infinite populations. The conditions usually given by corollaries or separate theorems are incorporated into one easily programmable expression.

The indices in the term corresponding to $s_i$ are represented by zeroes in binary expansion we denote by $s_i$. These zeroes can be partitioned into rightmost bracket and non-rightmost bracket indices by the zeroes in each of the two terms of the Boolean identity

$$s_i \equiv (s_i \lor \bar{r}_i) \land r_i .$$

Clearly, the zeroes in $r_i$ correspond to rightmost bracket indices exclusively, and the identity holds because, distributing the operators, we obtain

$$(r_i \land s_i) \lor (r_i \land \bar{r}_i) \equiv (r_i \land s_i) \equiv s_i ,$$

the last equivalence holding because $s_i \leq r_i$ is true by definition. Similarly, each $x \in J$ contains zeroes in positions corresponding to the subscripts associated with a particular $c_x^2$ candidate for the expansion of $E_{s_i}$. Again, the zeroes of $x$ can be subdivided according to rightmost and non-rightmost indices by the zeroes in the two terms of the
expression

\[ x = (x \lor \neg r^x) \land r^x . \]

If we now let

\[ W_x = r_x \lor \neg (r_x \lor s_i) \]

\[ = r_x \lor (\neg r_x \land \neg s_i) , \]

then, the position of the zeroes in \( W_x \) correspond to factors having finite population correction coefficients in the \( \sigma^2_x \) term entering the EMS.

The coefficient of \( \sigma^2_x \) is zero iff the zeroes of \( x \) do not contain the zeroes in \( s_i \). Hence, this condition is that if \( x \not\leq s_i \), then \( \sigma^2_x \) vanishes, or equivalently, we require \( s_i \lor x \) to be true and have binary representation of all \( l \)'s. Whenever the zeroes of \( s_i \) are contained in the zeroes of \( x \), i.e. \( x \leq s_i \), then the coefficient of \( \sigma^2_x \) in \( \text{EMS}^s_i \) contains the term

\[ x[w \circ R] , \]

which is the "number of times any one component whose type is specified by \( x \) enters into the complete sample used in the experiment".

It is known that if at least one of the excess subscripts of a rightmost bracket, i.e. at least one zero in the expansion of \( W_x \), has its sample range equal to its population range, then the contribution of that \( \sigma^2 \) vanishes from the expectation. This implies that a zero in \( W_x \) requires a zero in the corresponding position of the \( f \) vector previously defined. Thus, \( W_x \lor \neg f \) must be true and have a representation of all \( l \)'s if \( \sigma^2_x \) is...
is to enter the expected mean square. Finally, note that by construc-
tion of the vector c containing finite correction factors, those fac-
tors corresponding to indices of infinite population range were given
finite correction equal to unity. Consequently, the appropriate finite
correction coefficients are given by
\[ \mathcal{W} [\mathcal{W}_x \odot c] \,.

We now can give the complete equation for EMS, , for any s, ,
corresponding to any balanced complete structure and incorporating si-
multaneously all cases of "fixed", "semi-random", and "random" factors.

**Theorem 2.3** The expected value of the mean square of the line
corresponding to binary expansion s, has the form

\[ \text{EMS}_{s_i} = \sum_{x \in \mathcal{A}} \mathcal{W} [x \odot R] \mathcal{W} [\mathcal{W}_x \odot c] \sigma_x^2, \]

\[ \forall x \text{ such that} \]

\[ \left[ (W_x \lor \neg f) \lor (s_i \lor \neg x) \right], \]

where

\[ W_x \equiv r_x \lor \neg (r_x \lor s_i). \]

I. Extension to Balanced Incomplete Samples
from Balanced Complete Populations

It would seem highly desirable to adjust the preceding logical
formulation in such a way that it would apply in cases of data re-
sulting from such designs as latin squares, balanced incomplete
blocks, and the various lattice designs. Such data can be charac-
terized as special cases of balanced incomplete samples from a hy-
pothesized balanced complete population structure. Computationally,
the sampling procedures and restrictions leading to particular realizations of such designs are not, in general, totally described by the structures previously given. However, the orthogonality present does allow in each case simple computations and relatively simple expressions for expected mean squares, in terms of sample $\Sigma$'s.

For example, the A.O.V. for a latin square can be obtained from this logical formulation in two passes of data. The first specifies rows and treatments, ignoring columns, and the second specifies columns and treatments, ignoring rows. Proper combination of the two analysis yields the standard computations relative to a latin square design.

The expected mean squares, however, have relatively simple expressions in a much wider class of incomplete data. For any two responses in an observation structure, the observations are uniquely identifiable by the vector of levels of each factor, e.g.,

$$\mathbf{y}_x = (t_1, t_2, t_3, \ldots, t_n) \quad \text{and} \quad \mathbf{y}_{x'} = (t'_1, t'_2, \ldots, t'_n).$$

The $n$-dimensional vector $\mathbf{Z}$ with components

$$Z_i = \begin{cases} 1 & \text{if } t_i \neq t'_i \\ 0 & \text{if } t_i = t'_i \end{cases}$$

is useful in obtaining a simple logical formulation of the expectation of the product of any two responses in terms of population $\Sigma$'s. The general result as given by White is

$$E(\mathbf{y}_x \mathbf{y}_{x'}) = \sum_{x \in J} \Sigma_x.$$
where the sum is over all \( x \in \mathcal{J} \), such that \( x \subset Z \) is satisfied.

The expression above facilitates the computation of EMS's for a wider class of designs that those characterized by balanced samples of balanced populations. White (26) has given the appropriate results for such incomplete structures as those resulting from Finney partition, modified latin square, n-dimensional lattices, and the generalized incomplete block designs. A summary of this development may be found in the report by Zyskind, et al. (39). In the context of random experimental material only, the cap sigma quantities were, essentially, also used by Nelder (21).

J. A Comparison of the Development

In 1964 Nelder (21) published a two-part paper aimed at giving a unified development of the analysis of randomized experiments. Nelder's results form a development very similar to that on which this chapter was based. The formulation of the analysis as given in Kempthorne (18), Wilk (27), Zyskind (36), Throckmorton (24), and White (26) had chronological precedence. The main objective of this section is to examine those differences between these two developments which for balanced complete structures affect the performance of the actual computations. Because Nelder's terminology is distinct from that of the preceding sections, a few terms will be related in the following.

The class of experimental situations that Nelder considers can be characterized as those having:
1) Orthogonal block structure, and
2) Additive linear treatment structure.

An experiment is defined by Nelder as having orthogonal block structure if the identity decomposition of a typical response has a matrix representation that satisfies the conditions of Cochran's Theorem. Nelder implicitly defines linear treatment structures as those which can be specified by idempotent matrices similar to those used in specifying an orthogonal block structure. The additivity requirement of Nelder implies that there are no interactions between factors in the block structure and factors in the treatment structure.

The results of Zyskind and White provide the necessary equations for computation in experimental situations characterized by balanced samples from balanced complete populations. Their results are further useful in computing the expectations for certain incomplete designs, i.e., Latin squares, Balanced incomplete blocks, etc. Under this development no assumption is necessary of additivity of treatment effect and experimental material effect.

Nelder partitions the class of experiments having orthogonal block structure into simple block and non-simple block structures. He exhibits a complete randomization theory for simple block structures and outlines a method for computations. Under certain orthogonality assumptions, Nelder derives computational methods for experiments having non-simple orthogonal block structure.

The development given in this Chapter does not provide computational methods for the non-simple orthogonal block structure. Hence, Nelder's technique is more general. Nevertheless, we examine in this section some
important differences between the methods of computation for simple block structures.

Computationally, Nelder's method is based on a notation using two operators (\(\rightarrow\) denotes nesting, and \(\times\) denotes crossing) to describe simple block structures. A simple block structure is defined by "any formula involving the two operators, and suitable brackets to indicate order of combination, and the \(n_i\), one for each category, giving the number of units in that category".

These formulas are then used to establish the fundamental identity for yields, the covariance matrix of the null randomization distribution, the randomization model, and the expectations under randomization of the mean squares.

The formulas for the simple block structures of Nelder can readily be converted to the structure matrices defined in Section D. For example

\[
(B_1 \rightarrow B_2) \times B_3 = \begin{bmatrix}
0 & 1 & 1 \\
0 & 0 & 1 \\
1 & 1 & 0
\end{bmatrix}
\]

The advantages of the latter representation over that of Nelder are:

1. The structure matrix being formed by binary row vectors can be stored and manipulated by a digital computer.
2. The matrix structure as defined in Section D has a one-to-one correspondence with distinct experimental structures. By comparison, consider, for example, the representation of a four factor structure using Nelder's notation. There are twenty three possibilities given by:
\((^a \leftarrow ^a \times \bar{a}) \leftarrow T a\) (68)

\(v_a \leftarrow (\bar{a} \times \bar{a}) \leftarrow T a\) (68)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)

\((^a \leftarrow \bar{a}) \times \bar{a} \leftarrow T a\) (61)
20) \(((B_1 \rightarrow B_2) \times B_3) \rightarrow B_4\)
21) \(B_1 \rightarrow B_2 \rightarrow (B_3 \times B_4)\)
22) \((B_1 \rightarrow B_2 \rightarrow B_3) \times B_4\)
23) \(B_1 \rightarrow (B_2 \rightarrow B_3) \times B_4\).

Many of these are redundant; moreover, the structure diagrammatically and in structure matrix form represented by

\[
\begin{bmatrix}
1 & 1 & 1 & 1 \\
0 & 1 & 1 & 1 \\
1 & 0 & 1 & 1 \\
0 & 1 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0
\end{bmatrix}
\]

does not admit a one-line representation of the Nelder type. Hence, Nelder does not provide any analysis of this structure or many other like it because the computations, i.e., the identity decomposition randomization procedure, form of expected mean squares, etc., are based entirely on the notation previously discussed for simple block structures. If the above structure were specified in Nelder's notation by

\((B_1 \times B_2) \rightarrow B_4\), \quad \text{and} \quad \ B_2 \rightarrow B_3,\]

the algebra for computation would need to be modified considerably. Structures of more factors would require several lines of description.
The identity decomposition of a yield would follow a complex sequence in any imaginable revision of the recipe as given by Nelder. By comparison, the logical operators on the rows of the structure matrix defined in this chapter are basic computer operations and do the job easily and efficiently for all balanced complete structures.

In conclusion, it is desirable to indicate that the specifications for a computer program given by Nelder are inefficient for certain analyses. Nelder states the program requires three pieces of information to carry out the analysis. These are:

1) the block structure,
2) the design matrix \( N \),
3) the treatment model.

If one considers each in turn, it is found that:

1) The claim that the specified block structure defines the \( C_i \) matrices of the generalized Cochran Theorem is not valid since some structures do not admit specifications.

2) Nelder's design matrix \( N \) is of order \( n \times t \), where \( n \) is the number of experimental units and \( t \) is the number of treatments. A matrix of this size as input to a program is not to be considered feasible except for very small problems. Consider a split-split plot experiment with 5 whole plots treatments, 5 split plot treatments, and 5 split-split plot treatments. These results in 425 experimental units, even without repetitions.
Admitting interactions between the various treatments factors, the size of \( N \) required is \( 125 r \times 215 \), where \( r \) is the number of repetitions. Such experiments are not rare or even considered large. The methods previously discussed in this chapter have the significant advantage of being specifically computer-oriented and conservative in computer memory requirements.

3) Nelder's procedure requires construction of treatment structure matrices \( T_j \) from the treatment model, block structure matrices \( C_j \) from the block structure specification and the design matrix \( N \). Then if

\[
B_i = C_j N^T, \text{ where } T = \sum_j T_j,
\]

the resulting normal equation for what Nelder calls the \( j \)-th block strata would be

\[
(B_i^T B_i)(T_t) = B_i^T y.
\]

Here \( t \) is the vector of treatment parameters. The proposed solution involves the decomposition of \( B_i^T B_i \) into spectral form. The computations are quite complex except in the balanced case, requiring the eigenvalues and ranks of several matrices. In the general case, there also remains the issue of combining information on treatments from estimates obtained on the basis of the individual strata.

In the balanced complete case, the methods developed in this chapter provide directly all needed partial means and sums of squares. Attempts at implementing computational algorithms which apply to incomplete block designs have been made, for example, by Wilkinson (32). However, the implemented method does not follow the outline presented in Nelder (21).
III. THE NUMBER OF ADMISSIBLE SETS IN AN ARBITRARY STRUCTURE, AND THE NUMBER OF STRUCTURES CONTAINING AN ARBITRARY NUMBER OF FACTORS

A. Two Problems

Alternative ways of representing structures were given in Figure 2.1. The most convenient forms, for present purposes, are the structure diagrams of Throckmorton's, exemplified by Figure 2.1 (b) and defined in Definition 2.14, and the right-most bracket factor name convention, due to Zyskind, exemplified in Figure 2.1 (e) and defined in Definition 2.11. The work that follows relies heavily on the concepts of nesting, direct nesting, multiple nesting, crossing and graphs of response structures. These terms are used strictly in the context of the Definitions 2.5, 2.6, 2.7, 2.8, and 2.14 respectively.

Throckmorton (14) and in (19), exhibited all the structure diagrams for $N \leq 5$ where $N$ is the number of factors in the structure, excluding the general mean and error term. Gilbert (12) extended the sequence to the case where $N = 6$. The results established no general sequence which would allow determination of the number of structures for an arbitrary $N$. The method of the authors consisted of exhaustive enumeration, but this proved extremely tedious because the number of distinct structures increases very rapidly as $N$ increases. Section D, of this chapter, contains a sequence of inductive equations giving the number of distinct structures as a function of $N$. The induction is taken with respect to $q$, where $q$ is the number of factors directly nested (see Definition 2.6)
in factors other than the mean. For fixed values of \( q \), in the range \( q = 0, 1, 2, 3 \), a structure is given a characterization as an \( n \)-tuple of order \( n = 1, 1, 3, 8 \) respectively. The isomorphism between a structure and its \( n \)-tuple permits both the determination of the number of structures for fixed values of \( N \) and \( q \) and the simple enumeration of each such unique structure. The work contained in Section D provides a partial solution to the problem of determining the number of structures containing an arbitrary number of factors. The equations for \( q \geq 4 \), in following the present development, would have to be derived individually for each successive value of \( q \) because the general step in the induction was not obtained. Brief remarks with respect to generalization of the results are given in part 3 of Section D. However, the work described in Section D is intended primarily as an exposition of the complexities that arise in attaining a general solution for the given problem, and is directed towards a general recursive formulation programmable on a digital computer.

The inductive approach used in Section D was suggested by the results of exploring the problem of determining the number of admissible sets (see Definition 2.11) in an arbitrary \( N \)-factor structure. Section C of this chapter provides a combinatorial equation for the number of admissible sets in a structure as a function of parameters characterizing the nesting relationships of the factors in the structure. The notation, used in the development of the two problems, is given in Section B.
While the number of possible structures may be mainly of academic interest to statisticians, the number of admissible sets, corresponding to a given structure, has a practical and immediate application. In the design of general analysis of variance programs, it is frequently necessary to incorporate size parameters because of limitations on available computer memory. One such parameter is the number of admissible sets. For a given structure, the number of admissible sets is the same as the number of admissible model terms, admissible means, and admissible components. While for a small number of factors the exhaustive enumeration of admissible sets is feasible, for even moderately large numbers of factors the task is tedious and unnecessary. Chapter II describes a technique which enables a program to generate all the admissible sets. Hence, since the problem specification does not require the exhaustive enumeration, it is important to provide an algorithm to determine if a particular structure exceeds the maximum allowable limit of admissible sets.

B. Notational Conventions

Because the right-most bracket and the structure diagram are conventions heavily employed in this chapter, the reader is referred to Definitions 2.13 and 2.14 respectively. The relationship between the two notations is as follows. First, let the nodes of a structure diagram be labelled $\mu, A_1, A_2, \ldots, A_N, \epsilon$; where $\mu$ is the label corresponding to the node at the top of the diagram and $\epsilon$ is the label corresponding to the node at the bottom. The remaining $N$ labels may be assigned arbitrarily.
Each symbol corresponds to the right-most bracket symbol of a factor name, as given in Definition 2.13. Then to each symbol $A_i$, one may associate the set of symbols $V_i$ that correspond to those nodes attached to $A_i$ from above in the structure diagram and distinct from $\mu$. The set $V_i$ corresponding to $A_i$ contains the symbols in the factor name that are not in the right-most bracket. As a direct consequence, $V_i = \emptyset$ if $A_i$ is directly nested (see Definition 2.6) in the factor $\mu$. The above results in a label for each node of the diagram given by the following:

(3.1) \text{Set of Factor Names } = \mathcal{F}; \quad V_i(A_i), \ i = 1, 2, \ldots N; \ \mathcal{F}_j.

Note that the correspondence between the labels $A_i$ and the nodes of the diagram was made arbitrarily in the above. For subsequent purposes, it is convenient to require that the $A_i$ labels be assigned to the nodes in non-decreasing order according to the number of symbols in $V_i$. This convention for labelling the nodes facilitates the attainment of general formulas, since it will not be necessary to distinguish the structure corresponding to a labelling $A_2(A_1)$, $A_1A_2(A_2)$, and $A_2$ from that corresponding to $A_1$, $A_1(A_2)$, and $A_1A_2(A_3)$. The former are not admissible in our notation.

To distinguish between a set of symbols and the number of symbols in a set, $\eta$ will be used exclusively to denote the function that counts the symbols in its argument. For example, the monotonicity condition previously stated can be equivalently expressed:

(3.2) \quad \eta(V_i) \geq \eta(V_j) \quad \text{for all } i > j.
Whenever the complement of a set is used, the complementation is with respect to the universal set $U$, where $U = \{A_i; i = 1, 2, \ldots, N\}$. For example, the complement of $A_1$ is denoted by $\overline{A}_1$ and is equal to

$$\overline{A}_1 = \{A_i; i = 2, 3, \ldots, N\}.$$ 

**Definition 3.1**

A structure having diagrammatic representation with $N + 2$ nodes is called an $N$-factor structure.

### C. Number of Admissible Sets in an Arbitrary Structure

It is known that the structure containing $N$ crossed factors in addition to $\mu$ and $\varepsilon$ contains $2^N$ admissible sets. Therefore, for this structure, $M$, the number of admissible sets, can be expressed by

$$M = 2^N = \sum_{i=0}^{N} \binom{N}{i}.$$  (3.3)

In the expansion of the summation $\binom{N}{i}$ corresponds to the set for the mean, $\binom{N}{1}$ is the number of sets corresponding to main factorial effects, $\binom{N}{2}$ the number of sets corresponding to first order or 2-way interactions, and so on; $\binom{N}{N}$ corresponds to the one admissible set for the $N$-way interaction. The preceding enumeration is well known and is given, for example, in Kempthorne (18). By the convention given in (3.2), the preceding enumerates admissible sets for that unique $N$-factor structure having the property $\mathcal{Q}(V_N) = 0$.

Now, consider an $N$-factor structure with arbitrary crossing and nesting relationships among the factors. Let $C = \{A_i; i = 1, 2, \ldots, k\}$
be that subset of $U$ containing the symbols of factor names that are directly nested in $\mathcal{M}$ in the structure in question. Then, $A_i \in C$ implies $\gamma_i(V_i) = 0$, and $A_i \not\in C$ implies $\gamma_i(V_i) > 0$. These $k$ factors, and the interactions of all possible orders to which they give rise, contribute $2^k$ admissible sets. Consider, now, the remaining $N-k$ factors of the structure that are not elements of the set $C$ and consequently are contained in $\bar{C}$. If $A_i \in \bar{C}$ the number of admissible sets containing only $A_i$ and symbols from $C$ in the right-most bracket is given by:

$$\sum_{i=0}^{N-k} \binom{\gamma_i(C \cap \bar{V}_i)}{i} = 2^{\gamma(C \cap \bar{V}_i)}.$$ 

Therefore, the total number of admissible sets containing one symbol from $\bar{C}$ and an arbitrary number of symbols from $C$ is given by:

$$(3.4) \quad \sum_{j=1}^{N-k} \sum_{i=0}^{N-k} \binom{\gamma_i(C \cap \bar{V}_{k+j})}{j} = \sum_{j=1}^{N-k} 2^{\gamma_i(C \cap \bar{V}_{k+j})}.$$ 

To illustrate the preceding, consider the structure represented diagrammatically by:

```
This structure corresponds to a naming of factors as in
```
Then \( C = \{ A_1, A_2, A_3, A_4, A_5, A_6 \} \) and \( k = \eta(C) = 6 \).

Since \( V_7 = \{ A_1, A_2, A_3 \} \),
\( V_8 = \{ A_2, A_3, A_4 \} \), and
\( V_9 = \{ A_1, A_2, A_3, A_4, A_5, A_7, A_8 \} \),
it follows that
\[
\eta(C \cap V_7) = \eta(A_4 A_5, A_6) = 3,
\eta(C \cap V_8) = \eta(A_1, A_5, A_6) = 3, \text{ and}
\eta(C \cap V_9) = \eta(A_6) = 1.
\]

Hence, by (3.4)
\[
\frac{3}{2} 2^{\eta(C \cap V_{6+j})} = 2^3 + 2^3 + 2^1 = 18,
\]
and this enumerates the admissible sets given by

\[
\begin{array}{llll}
A_1 A_2 A_3 (A_7) & A_1 A_2 A_3 (A_4 A_5 A_6) & A_2 A_3 A_4 (A_1 A_5 A_8) \\
A_1 A_2 A_3 (A_4 A_7) & A_1 A_2 A_3 (A_4 A_5 A_6) & A_2 A_3 A_4 (A_1 A_6 A_8) \\
A_1 A_2 A_3 (A_5 A_7) & A_2 A_3 A_4 (A_8) & A_2 A_3 A_4 (A_5 A_6 A_8) \\
A_1 A_2 A_3 (A_6 A_7) & A_2 A_3 A_4 (A_1 A_8) & A_2 A_3 A_4 (A_1 A_5 A_6 A_8) \\
A_1 A_2 A_3 (A_4 A_5 A_7) & A_2 A_3 A_4 (A_2 A_6) & A_1 A_2 A_3 A_4 A_5 A_7 A_8 (A_9) \\
A_1 A_2 A_3 (A_4 A_6 A_7) & A_2 A_3 A_4 (A_6 A_8) & A_1 A_2 A_3 A_4 A_5 A_7 A_8 (A_6 A_9)
\end{array}
\]
Note that the sets are written according to the convention of the right-most bracket as given in Definition 2.13. Additionally, consider the admissible sets containing, in the right-most bracket, two or more symbols corresponding to nodes or factors directly nested in factors other than $\mu$. Equivalently, these are admissible sets containing two or more symbols from $\mathcal{C}$ in the right-most bracket.

Definition 2.8 gave a formal description of crossed factors. Now, for any pair of integers $(l_1, l_2)$ such that $l_1 < l_2$, it is convenient to define

$$\delta_{l_1, l_2} = 1 \quad \text{if factor } A_{l_1} \text{ is crossed with factor } A_{l_2},$$

$$= 0 \quad \text{otherwise.}$$

Then, the number of admissible sets containing two symbols from $\mathcal{C}$ in the right-most bracket can be expressed as

$$\sum_{l_1 = k+1}^{N-1} \sum_{l_2 = l_1 + 1}^{N} \delta_{l_1, l_2} \sum_{i=0}^{2} \left( \gamma(\mathcal{C} \cap \bigcup_{i=1}^{N} V_{l_i}) \right),$$

or, equivalently by

$$\sum_{l_1 = k+1}^{N-1} \sum_{l_2 = l_1 + 1}^{N} \delta_{l_1, l_2} \sum_{i=1}^{2} \gamma(\mathcal{C} \cap \bigcup_{i=1}^{N} V_{l_i}).$$

If $\gamma(\mathcal{C}) = N - k$ is greater than or equal to 3, the structure possibly yields admissible sets containing three symbols from $\mathcal{C}$ in their right-most bracket. The equation (3.6) can be extended by defining for all integers triples $(l_1, l_2, l_3)$ such that $l_1 < l_2 < l_3$, \[\text{?}\]
\[ \delta_{1,1, l_2, l_3} = 1 \text{ if factors } A_{11}, A_{12}, \text{and } A_{13} \text{ are crossed,} \]

\[ = 0 \text{ otherwise.} \]

Then the structure in question yields a number of admissible sets containing three symbols from \( \mathcal{C} \) in the right-most bracket, given by

\[ (3.7) \sum_{l_1 = k+1}^{N-2} \sum_{l_2 = l_1 + 1}^{N-1} \sum_{l_3 = l_2 + 1}^{N} \delta_{1,1, l_2, l_3} \gamma(C \cap \overline{\bigcup_{i=1}^{l_1} V_i}). \]

Generalizing the preceding, one obtains:

**Theorem 3.1**

Consider the structure containing factors given by

\[ \{ \mu; \{ V_i(A_i) \}, i = 1, 2, 3, \ldots, N; \mathcal{C} \}. \]

Let \( C = \{ A_i \}, \) such \( \gamma(V) = 0 \} \) and \( k = \gamma(C) \).

The number of admissible sets corresponding to the structure is given by:

\[ M = 2^\gamma(C) + \sum_{j=1}^{N-k} 2^\gamma(C \cap \overline{\bigcup_{k+j}^N V_i}) + \sum_{l_1 = k+1}^{N-2} \sum_{l_2 = l_1 + 1}^{N-1} \sum_{l_3 = l_2 + 1}^{N} \delta_{1,1, l_2, l_3} 2^\gamma(C \cap \overline{\bigcup_{i=1}^{l_1} V_i}) + \]

\[ (3.8) \sum_{l_1 = k+1}^{N-2} \sum_{l_2 = l_1 + 1}^{N-1} \sum_{l_3 = l_2 + 1}^{N} \delta_{1,1, l_2, l_3} 2^\gamma(C \cap \overline{\bigcup_{i=1}^{l_1} V_i}) + \]

\[ \ldots + \delta_{k+1, k+2, \ldots, N} 2^\gamma(C \cap \overline{\bigcup_{i=1}^{N} V_i}). \]

**Example 3.1**

Consider the structure diagrammatically represented by:
Then $C = \{A_1, A_2, A_3\}$, $\gamma(0) = 3$ and the non-zero tuples of equation (2.8) are:

$(4,6); (5,6); (5,7); (6,7)$ and $(5,6,7)$.

The number of admissible sets, by equation (3.8), is

$$M = 2^3 + 2\gamma(A_1^2 A_2^2) + 2\gamma(A_2^3 A_3^2) + 2\gamma(A_3^3 A_3) + 2\gamma(A_3^2 A_3) + 2\gamma(A_3 A_3) + 2\gamma(A_3 A_3) + 2\gamma(A_3 A_3) + 2\gamma(A_3 A_3) + 2\gamma(A_3 A_3) = 8 + 4 + 4 + 2 + 1 + 2 + 1 + 1 + 1 + 1 = 26.$$

It is of interest to consider if one can derive Theorem 3.1 by enumeration of non-admissible sets in a structure. The difficulties that ensue from approaching the complementary problems are strikingly similar to the problems that arise in Section D in enumerating the possible structures of $N$ factors. Indeed, the work contained in Section D was suggested by the following. Consider any factor $V_k(A_k)$, such that $\gamma(V_k) \neq 0$. This factor results in a number of sets being non-admissible which would be admissible in a completely crossed model. The sets formed by combining any one symbol of $V_k$ with $A_k$, for example, are non-admissible. There are $\binom{\gamma(V_k)}{1}$ such sets.
The sets formed by combining two symbols from $V_k$ with $A_k$ are also non-admissible, as are those sets obtained formally by combining one symbol from $V_k$ with one symbol not in $V_k$ or $A_k$, and with $A_k$.

There are then,

\[
\binom{\eta(V_k)}{2} + \binom{\eta(V_k)}{1} \left( \eta(V_k \cup A_k) \right)
\]

additional non-admissible sets. Generally, one can enumerate non-admissible sets resulting from the nesting $V_k(A_k)$, in which $m$ is the number of symbols in the non-admissible set. Note that every combination term must be conditioned by the requirements of sufficient symbols for selection, i.e. $\binom{x}{y} = 0$ if $x \leq y$.

**Lemma 3.1**

The number of sets not admissible due to a factor $V_k(A_k)$, in which $\eta(V_k) \neq 0$, and admissible with the completely crossed $N$ factor structure, is given by:

\[
E(V_k(A_k)) = \sum_{m=2}^{N} \sum_{i=1}^{m-1} \binom{\eta(V_k)}{m-1} \binom{\eta(V_k \cup A_k)}{i-1},
\]

where \(\binom{\eta(V_k)}{m-1} = 0\) if $\eta(V_k) < m-i$, and \(\binom{\eta(V_k \cup A_k)}{i-1} = 0\) if $\eta(V_k \cup A_k) < i - 1$. 
When the structure contains two or more factors directly nested in factors other than the mean, Lemma 3.1 can not be applied directly to each factor in order to obtain the number of non-admissible sets. For example, for a five factor structure with nodes labelled \( \{A, A_1, A_2, A_3, A_1(A_4), A_1(A_5)\} \), Lemma 3.1 applied to \( A_1(A_4) \) excludes as non-admissible the set consisting of \( A_1(A_4A_5) \). However, Lemma 3.1 applied to \( A_1(A_5) \) would exclude this same set again. Nevertheless, the equation of (3.9) of Lemma 3.1 is useful in motivating and suggesting subsequent generalizations.

Consider the recursive procedure where first, the number of sets that are non-admissible because of the \( N\)th factor is obtained. Next, one obtains the number of sets non-admissible because of the \( N-1\) factor, but not previously excluded, and so on until for some \( q \), \( \gamma(V_{N-q}) = 0 \), which terminates the process. Then, clearly if \( \gamma(V_{N-1}) = 0 \) the non-admissible sets are given by \( E(V(A_N)) \) of Lemma 3.1.

Suppose the structure under consideration is such that \( \gamma(V_{N-1}) \neq 0 \). All sets of factors, formally representable by combining symbols from \( V_{N-1} \) with the symbol \( A_{N-1} \), are non-admissible. Such sets were not excluded as non-admissible by \( E(V_{N}(A_{N})) \) of Lemma 3.1, because the latter comprised only sets containing the symbol \( A_N \), a symbol which by the convention given in (3.2) cannot appear in \( V_{N-1} \). Therefore, \( V_{N-1}(A_{N-1}) \) yields \( \binom{\gamma(V_{N-1})}{m-1} \) sets of \( m \) letters that are non-admissible in the given structure. A difficulty at this stage is exemplified by the following. Consider sets consisting of three symbols \( A_1A_jA_N \). If \( A_1 \in V_n \) and
$A_i \in V_{N-1}$, then the choices for $A_j$ exclude $A_N$, since $A_i A_{N-1} A_N$ is counted as non-admissible in $E(V_N(A_N))$. Conversely, if $A_i \in V_{N-1}$ and $A_i \notin V_N$, then $A_i A_{N-1} A_N$ is a non-admissible set and was not previously enumerated as such.

To systematize the generalization of this logical argument, with the aim of obtaining an expression for the number of non-admissible sets which are due to $V_{N-1}(A_{N-1})$ and are not counted in $E(V_N(A_N))$ of Lemma 3.1, the following sets are convenient. The symbols contained in $V_{N-1}$ can be partitioned into two disjoint sets $S_1(V_{N-1})$ and $S_2(V_{N-1})$ defined by:

$$S_1(V_{N-1}) = V_{N-1} \cap V_N,$$

$$S_2(V_{N-1}) = V_{N-1} \cap \overline{V_N}.$$  

Since $A_{N-1} \in V_N$ implies $V_{N-1} \subset V_N$, in this case $S_1(V_{N-1}) = \emptyset$ and $S_2(V_{N-1}) = V_{N-1}$. Hence, any set $A_i A_j A_{N-1}$ such that $A_i \in S_2(V_{N-1})$ and $A_j \notin (V_{N-1} \cup A_{N-1}) \cap A_N$, is a non-admissible set not counted by $E(V_N(A_N))$ of Lemma 3.1. Alternatively, if $A_{N-1} \notin V_N$, all sets of $m$ symbols combining one or more symbols of $V_{N-1}$ with symbols from $[V_{N-1} \setminus A_{N-1} \cap A_N]$ and with $A_{N-1}$ are non-admissible sets. Additionally, sets formally obtained combining symbols of $S_1(V_{N-1})$, one or more symbols from the set $\bigcap_{k=0}^{1} (V_{N-k} \cup A_{N-k})$, and $A_{N-1} A_N$, are non-admissible. In summary, the number of non-admissible sets due to the nested factor $V_{N-1}(A_{N-1})$, given that $E(V_N(A_N))$ were excluded by the nesting $V_N(A_N)$ is given by:
Case 1:
If $A_{N-1} \in V_N$, then
\[
\left( \eta(V_{N-1}) \right) + \sum_{i=1}^{m-2} \left( \eta(V_{N-1}) \right) \left( \frac{V_{N-1} \cup A_{N-1} \cap A_N}{m-i-1} \right)
\]
sets, consisting formally of $m$ symbols, are non-admissible.

Case 2:
If $A_{N-1} \not\in V_N$, then
\[
\left( \eta(V_{N-1}) \right) + \sum_{i=1}^{m-2} \left( \eta(V_{N-1}) \right) \left( \frac{V_{N-1} \cup A_{N-1} \cap A_N}{m-i-1} \right)
\]
\[
+ \sum_{i=1}^{m-2} \left( \eta(S_1(V_{N-1})) \right) \left( \frac{1}{k=0} \left( \frac{V_{N-k} \cup A_{N-k}}{m-i-1} \right) \right)
\]
sets, consisting formally of $m$ symbols, are non-admissible.

Lemma 3.2
The number of non-admissible sets resulting from an $N$ factor structure containing two factors $V_N(A_N)$ and $V_{N-1}(A_{N-1})$ directly nested in factors other than the mean is given by:

\[
(3.10) \quad E(V_N(A_N)) + \eta(A_{N-1} \cap V_N) + \sum_{m=2}^{N} \left( \eta(V_{N-1}) \right) +
\]
\[
+ \sum_{m=3}^{N-1} \left[ \sum_{i=1}^{m-2} \left( \eta(V_{N-1}) \right) \left( \frac{V_{N-1} \cup A_{N-1} \cap A_N}{m-i-1} \right) \right] +
\]
\[
+ \sum_{i=1}^{m-2} \left( \eta(S_1(V_{N-1})) \right) \left( \frac{1}{k=0} \left( \frac{V_{N-k} \cup A_{N-k}}{m-i-2} \right) \right),
\]
where $E(V_N(A_N))$ is defined as in Lemma 3.1.
Example 3.2

The eight factor structure, given by factors named

\[ \{ \mu, A_1, A_2, A_3, A_4, A_5, A_6, A_1A_2(A_7), A_1A_2A_3A_4(A_8) \} , \]

results in a partitioning of \( V_{N-1} = \{ A_1, A_2 \} \) into sets

\[ S_1(V_{N-1}) = \emptyset \quad \text{and} \quad S_2(V_{N-1}) = \{ A_1, A_2 \}. \]

By Lemma 3.1, \( \eta(V_N) = 4, \quad (V_N \cup A_N) = 3 \),

and consequently \( E(V_N(A_N)) = \sum_{m=2}^{3} \sum_{i=1}^{m-1} (\frac{4}{m-1}) \left( \frac{3}{i-1} \right) = 120 \).

By Lemma 3.2, \( \eta(A_{N-1} \cap V_N) = 1 \), and

\[ \sum_{m=2}^{3} \binom{2}{m-1} + \sum_{m=3}^{7} \left[ \sum_{i=1}^{m-2} \binom{2}{i} \left( \frac{4}{m-1-i} \right) \right] = 47. \]

Thus, 168 sets are non-admissible, and therefore the structure leads to \( 2^8 - 168 = 88 \) admissible sets.

Example 3.3

Consider the seven factor structure given by factor names

\[ \{ \mu, A_1, A_2, A_3, A_1(A_4), A_1A_4(A_5), A_1A_2(A_6), A_1A_2A_3A_4(A_7), \epsilon \} . \]

By Lemma 3.1, \( E[A_1, A_3, A_4(A_7)] = \sum_{m=2}^{7} \sum_{i=1}^{m-1} \binom{3}{i-1} \left( \frac{3}{m-1} \right) = 50. \)

By Lemma 3.2, the nesting \( A_1A_2(A_6) \), excludes additionally

\[ \sum_{m=2}^{7} \binom{2}{m-1} + \sum_{m=3}^{6} \left[ \sum_{i=1}^{m-2} \binom{2}{i} \left( \frac{3}{m-1-i} \right) + \sum_{i=1}^{m-2} \binom{1}{i} \left( \frac{1}{m-1-i} \right) \right] = 26 \]

sets.
This section concludes with an exposition of the problems of generalizing the present approach to structures containing an arbitrary number (greater than 2) of factors directly nested in factors other than $\mu$. Proceeding inductively let $\gamma(V_{N-k}) \neq 0$ for some $k \geq 2$. As before, those sets formally given by combining symbols of the $V_{N-k}$ with $A_{N-k}$ are non-admissible and not excluded in previous enumerations. The formal representations of such sets do not contain any of the symbols in the set $\{A_{N-1}, i = 0, 1, ..., k\}$, while previously enumerated non-admissible sets contained at least one symbol from the given set. The number of non-admissible sets that one can possibly form in this manner is

$$\sum_{m=2}^{N} \binom{\gamma(V_{N-k})}{m-1}.$$ 

It is now desirable to generalize the partitioning of the symbols of $V_{N-k}$. With respect to the symbols in non-rightmost brackets of factors whose enumeration has at the $k^{th}$ stage been considered, i.e., $\{V_{N-i}, i = 0, 1, ..., k-1\}$, any symbol $A_j \in V_{N-k}$ can occur in at most all $k$ non-rightmost brackets. The number of non-rightmost brackets in which $A_j$ occurs is $P(A_j) = \sum_{i=0}^{k-1} (A_j \cap V_{N-1})$. Grouping by distinct values of $P(A_j)$ defines the desired partitions of $V_{N-k}$. Then, $S_j(V_{N-k})$ is in general the set of symbols of $V_{N-k}$ such that $P(A_j) = j + 1$. Clearly the sets are exhaustive and disjoint since

$$V_{N-k} = \bigcup_{j=1}^{k+1} S_j(V_{N-k}).$$
As in Lemma 3.2, the number of sets consisting of m symbols containing one or more symbols of \( \bigcap_{i=0}^{k-1} A_{N-1} \) and \( A_{N-k} \) are not admissible. The resulting number is

\[
\sum_{i=1}^{m-2} \binom{V_{N-k}}{i} \binom{V_{N-k} \bigcap \bigcup_{l=0}^{k-1} A_{N-1}}{m-i-1}.
\]

Finally, the most difficult enumeration of non-admissible sets are sets containing three or more symbols \( A_{N-k} \), one or more symbols of the set \( \bigcup_{l=0}^{k-1} A_{N-l} \) and one or more symbols of \( V_{N-k} \). In general, \( \eta \left( \bigcup_{l=0}^{k-1} A_{N-1} \right) = k \), so that the partition of \( V_{N-k} \) results in k sets of eligible symbols.

The combinatorial generalization of Lemma 3.2 becomes increasingly complex.

D. On the Number of Structures Containing an Arbitrary Number of Factors

In the notation of Section B, the following develops an inductive set of functions, in one variable \( N \), giving the number of non-isomorphic structures which contain \( N \) factors plus a mean and error term. The induction is taken with respect to the number of factors which are directly nested in factors other than the mean, and this number is hereafter denoted by \( q \). Polynomial functions in \( N \) are obtained empirically for the case \( q = 2 \). However, these results do not yield a closed recursive formula for arbitrary \( q \). The procedure of decomposing the set of factors into disjoint exhaustive sets of like factors allows the extension of the induction one step at a time. A recursive programming
subroutine is currently being investigated as a means of providing a complete solution.

If \( q = 0 \), the \( N \) factors are crossed and only one such structure exists. This structure corresponds to what is often called the completely crossed \( N \)-factor structure. If \( q = 1 \), the structures have \( N-1 \) factors crossed and directly nested in the mean. The remaining factor may be nested, for example, in any one of the \( N-1 \) crossed factors, but this yields only one distinct structure since permutation of factor names do not alter the structure. In the context of Definition 1.1, the resulting structures are said to be isomorphic and therefore are not considered distinct. Thus, it is possible to view the \( N-1 \) nodes corresponding to the crossed factors as identical, at least for the purpose of identifying distinct structures. The determination of the number of structures will, in general, be based on this process of grouping \( N-1 \) nodes into sets of "like" nodes at each step of induction. In the case under consideration, \( q = 1 \), the partition is a trivial one, where one set, \( S_0 \), say, contains all \( N-1 \) nodes. Distinct structures are obtained by selecting a distinct number of nodes from \( S_0 \) as the nestors of \( A_N \), the factor directly nested in factors other than \( \mu_0 \). Trivially, then, there are \( N-1 \) non-isomorphic structures possible with \( N \) factors if only one factor is not directly nested in \( \mu \). For purposes of induction, the \( q = 1 \) case is given the following representation:

\[
(3.11) \quad S_0 = (A_j : j=1, \ldots, N-1)
\]
(3.12) \[ \gamma(S_0) = N - 1 \]

(3.13) \[ M_{N,1} = \sum_{i=1}^{N-1} \delta_i = N - 1 \text{ where } \delta_i = 1 \text{ if } i \leq \gamma(S_0) = N - 1, \]
\[ = 0 \text{ otherwise.} \]

\( M_{N,1} \) denotes the number of structures possible when \( q = 1 \) as a function of \( N \).

1. **Number of structures with two factors not directly nested in the mean**

If \( q = 2 \), that is, if the structures under consideration contain two nested factors not directly nested in the mean, it is possible to apply equation 3.13 with one less factor. Clearly, the result for \( q = 1 \) then gives \( N - 2 \) possibilities to the first nesting. For convenience, we will refer to these as **partial structures** since they have not been completed by the insertion of the factor \( A_N \) and the error term. Each of the \( N - 1 \) partial structures defines a partition of the \( N - 1 \) nodes into disjoint exhaustive sets defined by:

\[ S_0 = \text{nodes corresponding to factors directly nested in } \mu \text{ and not the nestors of the first nested factor;} \]

(3.14) \[ S_1 = \text{nodes corresponding to factors that are directly nested in } \mu \text{ and are nestors of the first nested factor;} \]

\[ S_2 = \text{the node corresponding to the factor, } A_{N-1}. \]

The number of symbols contained by each set in (3.14), in the notation of Section B, is given by:
\( \eta(S_0) = N - \eta(V_{N-1}) - 2, \)

(3.15) \( \eta(S_1) = \eta(V_{N-1}), \)

\( \eta(S_2) = 1. \)

For example, if \( N = 8, \) one of the six possible partial structures is:

\[
\begin{align*}
A_1 &- A_2 &- A_3 &- A_4 &- A_5 \\
A_7 &- A_6 &- A_5 &- A_4 &- A_3 \\
A_2 &- A_3 &- A_4 &- A_5 &- A_6
\end{align*}
\]

and,

\[
\begin{align*}
S_0 &= (A_4, A_5, A_6), \\
S_1 &= (A_1, A_2, A_3), \\
S_2 &= (A_7).
\end{align*}
\]

In completing the above structure, \( \eta(V_8) \) must be greater than or equal to three. The importance of the definition of the sets is that indetermining the number of ways of completing the structure, a permutation of nodes in the same set results only in isomorphic structures. That is, nodes in \( S_0 \) are essentially identical one to the other, while the same is true of nodes in \( S_1. \)

Now, in general, \( \eta(V_{N-1}) \) can range over values \( i = 1, 2, \ldots, N-2. \)
The condition imposed by equation 3.2 of Section B requires that 
\( \gamma(N_{N-1}) \leq \gamma(N_N) \). Therefore, \( \gamma(N_{N_{N-1}}) = j \) can have the values \( j = i, i+1, \ldots, N-1 \). For a fixed value of the pair \((\gamma(N_{N_{N-1}}), \gamma(N_N))\),

that is, for fixed values of \((i, j)\), it is possible to enumerate the number of non-isomorphic structures that are possible. For clarity, this can be paraphrased by saying that what follows gives the number of structures of \( N \) factors, such that each structure contains two factors directly nested in factors other than the mean, and such that the first is nested in \( i \) factors and the second in \( j \) factors. It is convenient to divide the structures into three classes.

**Class 1:** This class is characterized by \( i = j \). In this case, the node corresponding to \( A_{N} \) cannot be attached diagrammatically to the node corresponding to \( A_{N-1} \). The nodes eligible for connection with \( A_{N} \) are those contained in the sets \( S_0 \) and \( S_1 \). Distinct structures are obtained if and only if the proportions to the total \( i \) contributed by each set are varied. It is possible to choose the \( i \) nodes by selecting \( j' = 0, 1, \ldots, i \) from \( S_0 \), and \( i-j' \) from \( S_1 \), provided \( j' \leq \gamma(S_0) = N-i-2 \) and \( i-j' \leq \gamma(S_1) = i \). Then,

\[
\sum_{i=1}^{N-2} \sum_{j'=0}^{1} \delta_{j'}
\]

where \( \delta_{j'} = 1 \) if \( j' \leq N-i-2 \),

\( = 0 \) otherwise,
is the total number of such structures possible.

**Class 2:** This class is characterized by \( i < j \), with \( A_{N-1} \) and \( A_N \) not connected to each other diagrammatically. Again, \( S_0 \) and \( S_1 \) are the eligible contributors of connections for \( A_N \). Distinct structures are obtained if, and only if, the proportions to the total \( j \) contributed by each eligible set is varied. In this case \( j = i+1, \ldots, N-2 \). The number selected from \( S_0, j' \), can have values from 0 to \( j \). Hence,

\[
\sum_{i=1}^{N-2} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{i} \delta_{j'}^{i} = 1 \quad \text{if} \quad j' \leq N-i-2,
\]

\[
= 0 \quad \text{otherwise},
\]

is the total number of such structures possible.

**Class 3:** This class is characterized by \( i < j \) with factors \( A_{N-1}, A_N \) connected diagrammatically. Clearly, in the notation of Section B, \( A_{N-1} \in V_N \) implies that \( A_j \in V_N \) for all \( A_j \in S_1 \). Equivalently, any factor in which \( A_{N-1} \) is nested is also a nestor of \( A_N \). Hence, \( j \) can have values \( j = j+1, i+2, \ldots, N-1 \). The only eligible set is \( S_0 \) and it is required that for a fixed \( j \), \( S_0 \) contribute \( j-i-1 \) nodes to be connected with \( A_N \). Thus,

\[
\sum_{i=1}^{N-2} \sum_{j=i+1}^{N-1} \delta_{j'}^{i} = 1 \quad \text{if} \quad j' \leq N-i-2,
\]

\[
= 0 \quad \text{otherwise},
\]
where \( \delta_{j}^{\text{xx}} \) = 0 if \( j-i-l \leq N-i-2 \), and
\[ = 0 \quad \text{otherwise,} \]
is the total number of such structures possible.

The three classes are summarized in the following Lemma.

**Lemma 3.3**

The number of non-isomorphic structures possible with \( N \) factors, in which two factors of the structure are nested directly in factors other than the mean, is given by the expression:

\[
(3.16) \quad M_{N,2} = \sum_{i=1}^{N-2} \left\{ \sum_{j=0}^{i} \delta_{j}^{i} + \sum_{j=i+1}^{N-2} \sum_{j'=0}^{i} \delta_{j}^{j'} + \sum_{j=i+1}^{N-1} \right\},
\]

where \( \delta_{j}^{i} = 1 \) if \( j' \leq N-i-2 \),
\( = 0 \quad \text{otherwise,} \)
and \( \delta_{j}^{j'} = 1 \) if \( j' \leq N-i-2 \) and \( j-j' \leq i \)
\( = 0 \quad \text{otherwise.} \)

In expression 3.16, note that
\( \delta_{j}^{i} = 1 \) for structures in which the two factors, not directly nested in the mean, are nested in the same number of factors,
\( \delta_{j}^{j'} = 1 \) for structures in which the two factors, not directly nested in the mean, are nested in a distinct number of
factors, but the second nested factor is not nested in the first nested factor. The last term enumerates the number of structures possible in which the two factors not directly nested in the mean are themselves nested one in the other.

Although expression (3.16) is convenient in its present form for purposes of extending the enumeration to the case \( q = 3 \), it is possible to simplify the expression for ease in evaluation. Clearly, (3.16) may be equivalently written as:

\[
M_{N, 2} = \sum_{i=1}^{N-2} \left\{ \min(i, N-i-2) \left( \sum_{j=0}^{N-2} \sum_{i' = 1}^{\min(j, N-1)} \frac{1}{j' + \max(0, j-1)} + \sum_{j=1}^{N-1} \frac{1}{j+1} \right) \right\}.
\]

Then,

\[
M_{N, 2} = \sum_{i=1}^{N-2} \left\{ \sum_{j=1}^{N-2} \frac{\min(i, N-i-2)}{\min(j, N-1)} + \sum_{j=1}^{N-1} \frac{1}{j+1} \right\},
\]

or,

\[(3.17) \quad M_{N, 2} = \sum_{i=1}^{N-2} \sum_{j=1}^{N-1} \frac{\min(i, N-i-2)}{\min(j, N-i-1)} \cdot\]

Now, while each structure possible with \( q = 1 \) is characterized by a singleton (i), where i denotes the number of nestors of the nested factor, equation (3.17) gives a characterization of the structures possible when \( q = 2 \) in terms of a 3-tuple (i, j, j'). This result is summarized by:
Lemma 3.4

The number of non-isomorphic structures possible with N factors, in which two factors of the structure are nested in factors other than the mean, can be computed by evaluating the expression:

$$M_{N,2} = \sum_{i=1}^{N-2} \sum_{j=1}^{N-1} \sum_{j'=\max(0,j-i-1)}^{\min(j,N-i-2)}.$$

Corollary 3.1 To each N-factor structure containing two factors not directly nested in the mean, there corresponds a triple \((i, j, j')\), where \(i\) denotes the number of nestors of the first nested factor, \(j\) is the number of nestors of the second nested factor, and \(j'\) is the number of nestors of the second nested factor not nestors of first. To each integer \((i, j, j')\), such that \(1 \leq i \leq N-2\), \(i \leq j \leq N-1\) and \(\max(0, j-i-1) \leq j' \leq \min(j',N-i-2)\), there corresponds a N-factor structure. The correspondence is 1:1 onto and a structure is characterized equivalently by its corresponding triple.

The expression in Lemmas 3.3 and 3.4 were programmed and the program appears in the Appendix. The results of that corresponding to Lemma 3.4 for values of \(N \leq 30\) appear in Table 3.1. The first two forward finite differences were calculated in order to attempt to understand the nature of the relationship between \(N\) and \(M_{N,2}\). It was observed that the second difference is predictable and leads to the following results. From Table 3.1,
\[ M_{N,2} = \sum_{i=1}^{N-2} \Delta_i \] where \( \Delta_i \) denotes the \( i \)-th first forward finite difference. But \( \Delta'_i = 2 + \sum_{j=1}^{i-1} \Delta_j \) where \( \Delta_j \) denotes the \( j \)-th second forward finite difference.

If,
\[
\Delta_j'' = \begin{cases} 
2 + \frac{i}{2} & \text{if } j \text{ is even} \\
2 + \frac{i+1}{2} & \text{if } j \text{ is odd,}
\end{cases}
\]

then,
\[
M_{N,2} = \sum_{i=1}^{N-2} \left\{ 2 + \sum_{j=1}^{i-1} \Delta_j'' \right\}
\]
\[
= \sum_{i=1}^{N-2} \left\{ 2 + \sum_{j=1}^{i-1} \left( 2 + \frac{i}{2} \right) + \sum_{j=1}^{i-1} \left( 2 + \frac{i+1}{2} \right) \right\} \]
\[
= \sum_{i=1}^{N-2} 2 + \sum_{i=1}^{N-2} \sum_{j=1}^{i-1} \frac{i}{2} + \sum_{i=1}^{N-2} \sum_{j=1}^{i-1} \frac{i+1}{2} \]
\[
= \sum_{i=1}^{N-2} (2i + 1/4(i-1) (i)) + \sum_{i=1}^{N-2} \frac{i-1}{2} \]
\[
= \sum_{i=1}^{N-2} \left( 1/4 i^2 + 7/4 i \right) + \frac{1}{2} \sum_{i=1}^{N-2} \frac{i-1}{2}
\]
\[
= \frac{1}{24} \left( 2N^3 + 12N^2 - 50N + 36 \right) + \frac{1}{2} \sum_{i=1}^{N-2} \sum_{\substack{j=1\ \text{j odd}}}^{i-1} l.
\]

But, \[
\sum_{\substack{j=1\ \text{j odd}}}^{i-1} \frac{1}{2} = \begin{cases} 
\frac{1}{2} & \text{if } i \text{ is even,} \\
\frac{i-1}{2} & \text{if } i \text{ is odd}
\end{cases}
\]

and
\[
\sum_{i=1}^{N-2} \frac{i-1}{2} + \sum_{i=1}^{N-2} \frac{i}{2} = \sum_{i=1}^{N-2} \frac{i-1}{2} - \frac{1}{2} \sum_{i=1}^{N-2} 1;
\]

therefore,
\[
\left\{ \begin{array}{ll}
\frac{1}{8} (N^2 - 3N + 2) - 1/8 (N-2) & \text{if } N \text{ even} \\
1/8 (N^2 - 3N + 2) - 1/8 (N-1) & \text{if } N \text{ odd}
\end{array} \right.
\]

Thus,
\[
M_{n,2} = \left\{ \begin{array}{ll}
1/24 \left( 2N^3 + 15N^2 - 62N + 48 \right) & \text{if } N \text{ is even} \\
1/24 \left( 2N^3 + 15N^2 - 62N + 45 \right) & \text{if } N \text{ is odd}
\end{array} \right.
\]

We now have the result that for \( q = 1 \) and 2, the total numbers of possible non-isomorphic structures are given by polynomial equations of degrees 0, 1 and 3 respectively. The degree of the polynomial thusfar coincides with the \( n \)-tuple necessary to characterize the structure. The case for \( q = 3 \) was studied in hopes that the result would generalize.
Table 3.1. Number of structures containing two factors not directly nested in the mean, and the first two forward differences of the sequence.

<table>
<thead>
<tr>
<th>N</th>
<th>N° of factors</th>
<th>P(N)</th>
<th>N° of structures</th>
<th>Δ_j</th>
<th>Δ_j'</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>15</td>
<td>27</td>
<td>8</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>43</td>
<td>64</td>
<td>12</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>90</td>
<td></td>
<td>21</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>122</td>
<td>160</td>
<td>26</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>205</td>
<td></td>
<td>45</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>257</td>
<td>317</td>
<td>52</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>385</td>
<td></td>
<td>60</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>462</td>
<td>548</td>
<td>68</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>644</td>
<td>750</td>
<td>77</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>995</td>
<td></td>
<td>86</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1135</td>
<td>1462</td>
<td>96</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1287</td>
<td></td>
<td>106</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1452</td>
<td>1630</td>
<td>117</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1822</td>
<td>2028</td>
<td>128</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2249</td>
<td></td>
<td>140</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2485</td>
<td>2737</td>
<td>152</td>
<td>12</td>
<td></td>
</tr>
</tbody>
</table>
Table 3.2 Number of structures containing N factors, of which two are not directly nested in the mean. (I,J) is the respective number of factors in which the two factors are nested.

<table>
<thead>
<tr>
<th>N=3</th>
<th>I</th>
<th>J</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=4</th>
<th>I</th>
<th>J</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=5</th>
<th>I</th>
<th>J</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=6</th>
<th>I</th>
<th>J</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td></td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>6</td>
<td>8</td>
</tr>
</tbody>
</table>

27
Table 3.2. (continued)

<table>
<thead>
<tr>
<th>N=7</th>
<th></th>
<th>J</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>3</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>6</td>
<td>10</td>
<td>11</td>
<td>9</td>
<td>5</td>
<td>43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=8</th>
<th></th>
<th>J</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>14</td>
<td>11</td>
<td>5</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>N=9</th>
<th></th>
<th>J</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>-</td>
<td>-</td>
<td>4</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>2</td>
<td>6</td>
<td>11</td>
<td>16</td>
<td>18</td>
<td>17</td>
<td>13</td>
</tr>
</tbody>
</table>
The results of evaluating the expression of Lemma 3.3 are given in Table 3.2 for values of $N \leq 9$. The total number of non-isomorphic structures for each $N$ appears in the lower right hand corner of each sub-table. The individual entries in each square, for example, $N = 9, I = 3, J = 4$, indicate that there are five non-isomorphic structures when one factor is nested in 3 and the other in 4. These five structures are:

\begin{align*}
A_1^{I}A_2^{J}(A_3) & \quad A_1^{I}A_2^{J}(A_8) \\
A_1^{I}A_2^{J}(A_9) & \quad A_1^{I}A_2^{J}(A_9)
\end{align*}
2. **Number of structures with three factors not directly nested in the mean**

If \( q = 3 \), the condition given in (3.2) for \( q = 2 \) now becomes

\[
\eta(N-3) \leq \eta(N-1) \leq \eta(N).
\]

Therefore, if \( \eta(N-2) = i \), then \( i \) ranges from 1 through \( N-3 \); \( \eta(N-1) = j \) ranges from \( i \) through \( N-2 \), and \( \eta(N) = k \) can have values from \( j \) through \( N-1 \). By Lemma 2.3, the total number of partial structures containing two factors not directly nested in the mean is:

\[
\sum_{i=1}^{N-3} \left\{ \sum_{j'=0}^{\min(i,N-3)} 1 + \sum_{j'=\max(0,i-1)}^{\min(j,N-3)} 1 + \sum_{j'=i+1}^{N-2} 1 \right\}.
\]

Each partial structure defines a partitioning of the first \( N-1 \) nodes into disjoint exhaustive sets of like nodes. As before, the partitioning is defined so that permutations of nodes within each set results in isomorphic structures, while permutations of nodes between sets result in non-isomorphic structures. For the case when \( q = 3 \), the sets are defined by:

\[
S_0 \equiv \text{the set of nodes corresponding to direct nestors of } \mu, \text{ and not attached diagrammatically to } A_{N-2} \text{ or } A_{N-1};
\]

\[
S_{1,N-2} \equiv \text{the set of nodes corresponding to direct nestors of } \mu, \text{ and attached diagrammatically to } A_{N-2}, \text{ but not to } A_{N-1};
\]

\[
S_{1,N-1} \equiv \text{the set of nodes corresponding to factors direct nestors of } \mu, \text{ attached diagrammatically to } A_{N-1}, \text{ but}
\]
not to $A_{N-2}$.

$S_2$ = the set of nodes corresponding to factor direct nestors of $\mu$, attached diagrammatically to $A_{N-1}$ and $A_{N-2}$.

$S_{3,N-2}$ = the set containing the node corresponding to the first factor not directly nested in $\mu$, namely $A_{N-2}$.

$S_{3,N-1}$ = the set containing the node corresponding to the second factor not directly nested in $\mu$, namely $A_{N-1}$.

By Corollary 3.1, each partial structure is characterized by the parameters $(i, j, j')$ defined in Section B. It is desirable to represent the number of nodes contained by each set in (3.19) as a function of these parameters. Consider the set $S_0$. Clearly, there are $N-3$ factors directly nested in the mean $\mu$. From these, $i$ are direct nestors of $A_{N-2}$, and an additional $j'$ are direct nestors of $A_{N-1}$. Consequently, $\gamma(S_0) = N-3-i-j'$. Consider now the definition of $S_{1,N-2}$. If $A_{N-2}$ is not attached diagrammatically to $A_{N-1}$, then $S_{1,N-2}$ contains $i-j+j'$ nodes. Alternatively, if $A_{N-2}$ is a direct nestor of $A_{N-1}$, then $i-j+j' < 0$ and $S_{1,N-2} = \emptyset$. Consequently, for every partial structure,

$$\gamma(S_{1,N-2}) = \max (0, i-j+j').$$

In the special case when $i = j$, the nodes in the set $S_{1,N-2}$ and those of $S_{1,N-1}$ are identical for purposes of defining non-isomorphic structures. That is, when $i = j$ permutation of nodes between these two sets leaves the structure unchanged. Note that $\gamma(S_{1,N-1}) = j'$ in all partial structures.
Let $\delta^p_q$ be a zero-one random variable defined by:

$$
\delta^p_q = \begin{cases} 
1 & \text{if } p > q, \\
0 & \text{if } p \leq q.
\end{cases}
$$

The special situation presented by the case $i = j$ is then resolved by setting $\gamma(S_{1,N-2}) = \max(0, i-j+j') + j' \delta^j_{j-1}$, and $\gamma(S_{1,N-1}) = j' \delta^j_i$.

In the case $i = j$, $\gamma(S_{1,N-2}) = \max(0, i-j+j') + j$ and $\gamma(S_{1,N-1}) = 0$, while $i < j$ results in the number of nodes contained by these two sets given by:

$$
\gamma(S_{1,N-2}) = \max(0, i-j+j'), \quad \gamma(S_{1,N-1}) = j'.
$$

There remains to be considered the number of elements in the set $S_2$. This is given by $\eta(S_2) = j-j'$ in structures where $A_{N-2}$ is not a direct nestor of $A_{N-1}$, and by $i$ in those structures where $A_{N-2}$ is a direct nestor of $A_{N-1}$. In summary, the number of elements in each of the above sets are summarized for all partial structures by:

$$
\eta(S_0) = N-3-i-j',
$$

$$
\eta(S_{1,N-2}) = \max(0, i-j+j') + j' \delta^j_{j-1},
$$

$$
\eta(S_{1,N-1}) = j' \delta^j_i,
$$

$$
(3.20) \quad \eta(S_2) = j-j' - \delta^{j-i}_{j'},
$$

$$
\gamma(S_{3,N-2}) = 1,
$$

$$
\gamma(S_{3,N-1}) = 1.
$$
Using the expressions in (3.18), (3.19) and (3.20), the determination of the number of structures when \( q = 3 \) proceeds in the following logical sequence:

**General Procedure**

**Step 1:** Partition each of the three classes of structures, defined for structures with \( q = 2 \), into mutually disjoint and exhaustive classes which result from the addition of a third factor not directly nested in the mean.

**Step 2:** Determine which of the sets defined in (3.19) are eligible contributors, to the \( k \) symbols of the non-rightmost bracket of the added factors.

**Step 3:** Obtain and evaluate an expression giving the possible ways of varying the proportions contributed by each of the eligible sets as determined in **Step 2**.

The result of **Steps 1 and 2** are summarized in Tables 3.3, 3.4 and 3.5. That the classes of structures defined in the tables are disjoint and exhaustive is clear from the definitions of each class and is not elaborated upon in the interest of brevity.
Table 3.3. Breakdown of Class 1 structure \((i=j)\) into classes of structures for the case \(q = 3\), and the eligible contributors from each partial structure

<table>
<thead>
<tr>
<th>Structure Class</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i=j=k)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A_{N-2} \notin V_N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(A_{N-1} \notin V_N)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(S_0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(S_{1,N-2})</td>
<td>1</td>
<td>1</td>
<td>(X)</td>
<td>(X)</td>
</tr>
<tr>
<td>(S_{1,N-1})</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>(X)</td>
</tr>
<tr>
<td>(S_2)</td>
<td>1</td>
<td>(\cdot)</td>
<td>1</td>
<td>(X)</td>
</tr>
<tr>
<td>(S_{3,N-2})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>(X)</td>
</tr>
<tr>
<td>(S_{3,N-1})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>(X)</td>
</tr>
</tbody>
</table>

Legend: 1 = eligible contributor to last nesting.  
0 = not eligible as contributor to last nesting.  
\(X\) = included in last nesting by conditions of the class.
Table 3.4. Breakdown of Class 2 structures \((i < j, A_{N-2} \notin V_{N-1})\) into classes of structures for the case \(q = 3\), and the eligible contributors from each partial structure

<table>
<thead>
<tr>
<th>Structure</th>
<th>2.1</th>
<th>2.2</th>
<th>2.3</th>
<th>2.4</th>
<th>2.5</th>
<th>2.6</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i &lt; j = k)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
</tr>
<tr>
<td>(i &lt; j &lt; k)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
<td>1 (\notin V_N)</td>
</tr>
<tr>
<td>(s_0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(s_{1,N-2})</td>
<td>1</td>
<td>X</td>
<td>1</td>
<td>X</td>
<td>1</td>
<td>X</td>
</tr>
<tr>
<td>(s_{1,N-1})</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>(s_2)</td>
<td>1</td>
<td>X</td>
<td>1</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>(s_{3,N-2})</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>0</td>
<td>X</td>
</tr>
<tr>
<td>(s_{3,N-1})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Legend:  
1 = eligible contributor to last nesting.  
0 = not eligible as contributor to last nesting.  
X = included in last nesting by conditions of the class.
Table 3.5. Breakdown of Class 3 structures \((i < j, A_{N-2} \notin V_{N-1})\) into classes of structures for the case \(q = 3\), and the eligible contributors from each partial structure.

<table>
<thead>
<tr>
<th>Structure Class</th>
<th>Class 3.1</th>
<th>Class 3.2</th>
<th>Class 3.3</th>
<th>Class 3.4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i &lt; j = k)</td>
<td>(A_{N-2} \in V_N)</td>
<td>(A_{N-2} \notin V_N)</td>
<td>(A_{N-2} \notin V_N)</td>
<td>(A_{N-2} \notin V_N)</td>
</tr>
<tr>
<td>(S_0)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(S_{1,N-2})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>(S_{1,N-1})</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>X</td>
</tr>
<tr>
<td>(S_2)</td>
<td>X</td>
<td>1</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>(S_{3,N-2})</td>
<td>X</td>
<td>0</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>(S_{3,N-1})</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>X</td>
</tr>
</tbody>
</table>

Legend: 1 = eligible contributor to last nesting.  
0 = not eligible as contributor to last nesting.  
X = included in last nesting by conditions of the class.
Corollary 3.1 states that the structures with $q = 2$ have a representation in terms of a tuple $(i, j, j')$. For the case $q = 3$, it will be required to use then additional parameters, $k, k', k'', k''', k^IV$. The parameter $k$ has been defined as the number of factors other than $\mu$ in which $A_N$ is nested. Then, $k', k'', k''', k^IV$ will be used in the following to represent the elements of $S_0, S_1, N-2, S_2, N-1$, and $S_2'$ respectively, in which $A_N$ is nested.

By referring to Table 3.3, it can be seen that the class of structures containing two factors not directly nested in the mean, and characterized by $i = j$, can be decomposed into four distinct classes of structures with the inclusion of a third factor not directly nested in the mean. For example, if $N = 4$ and $i = j = 1$ the enumeration previously given yielded two distinct partial structures of class 1:

![Diagram](image)

If the structures are completed by the addition of a third nested factor and the error term, four classes of structures labelled 1.1, 1.2, 1.3 and 1.4 can result. For example, the left-hand structure diagrammed above left can be completed as follows, illustrating the four basic classes:
The four subclasses as defined, are disjoint and exhaust the possible completions of Class 1 partial structures.

The first of these subclasses of structures, Type 1.1, is characterized by \( i = j = k \), where \( i, j, k \) are the number of factors in which the three factors, not directly nested in mean, are nested. From the expressions in (3.19) and Table 3.3, the number of Class 1.1 structures possible for a given value of \((N, i, j')\) may be expressed by:
where \( \delta_{k',k''} = 1 \) if \( k' \leq h(S_0) = N-3-i-j' \)
and \( k'' \leq h(S_1,N-2) = 2_j' \)
and \( i-k'-k'' \leq h(S_2) = i-j' \)

= 0 otherwise.

However, some of the resulting structures will be isomorphic to those resulting from the completion of a partial structure with parameters \( (N, i, j') \). For example, the two partial structures previously given with characteristics

\[
N = 4, \ i = 1, \ j' = 1 \quad N = 4, \ i = 1, \ j' = 0
\]

\[
\begin{align*}
A_1 & \quad A_2 \\
A_3 & \quad A_4 
\end{align*}
\quad
\begin{align*}
A_1 & \quad A_2 \\
A_3 & \quad A_4
\end{align*}
\]

can result in complete structures

\[
N = 4, \ i = 1, \ j' = 1, \ k' = 0, \ k'' = 1 \quad N = 4, \ i = 1, \ j' = 0, \ k' = 1, \ k'' = 0
\]

\[
\begin{align*}
A_1 & \quad A_2 \\
A_3 & \quad A_4 \\
A_5 & \\
\end{align*}
\quad
\begin{align*}
A_1 & \quad A_2 \\
A_3 & \quad A_4 \\
A_5 & \\
\end{align*}
\]
which are isomorphic. This difficulty is avoided by modifying (3.21) and eliminating the enumeration of structures such that

\[ j' < k' \]

where \( j' \) is the number of factors in which \( A_4 \) is nested but \( A_3 \) is not, and \( k' \) is the number of factors in which \( A_5 \) is nested but \( A_3 \) and \( A_4 \) are not. The structure on the right then becomes non-admissible. The following Lemma summarizes the above.

**Lemma 3.5**

The number of non-isomorphic structures possible with \( N \)-factors having the properties:

1) three factors are directly nested in factors other than the mean,

2) each of the factors, not directly nested in the mean, is nested in the same number of factors,

is given by the expression:

\[
\sum_{i=1}^{N-3} \min \left( i, N-i-3 \right) \sum_{j'=0}^{i-3} \min \left( i, N-3-i-j' \right) \sum_{k'=0}^{i-3} \min \left( i-k', 2j' \right)
\]

\[
\min \left( i-k'-k'', i-j' \right)
\]

\[
\sum_{k''=\max \left( 0, i+2j'+3-N \right)} \delta_{j', k''}
\]

where \( \delta_{j', k'} = 1 \) if \( j' \geq k' \),

\( = 0 \) otherwise.

The computed results for some values of \( N \) in expression (3.22) are given in Table 3.6.
For the class of structures denoted by 1.2 in Table 3.3, the number of structures as a function of $N$ is given by the expression:

$$
(3.23) \sum_{i=1}^{N-3} \sum_{j=0}^{\min(i,N-i-3)} \sum_{k=i+1}^{N-3} \sum_{k'=0}^{k} \sum_{k''=0}^{k-k'} \delta_{k',k''}
$$

where $\delta_{k',k''} = 1$ if $k' \leq \gamma(S_0) = N-3-i-j'$,

and $k'' \leq \gamma(S_1,N-2) = 2j'$,

and $k-k'-k'' \leq \gamma(S_2) = i-j'$,

$= 0$ otherwise.

**Lemma 3.6**

The number of structures possible with $N$ factors having the properties

1) three factors are not directly nested in the mean,

2) the first two such nested factors are nested in the same number ($i$) of factors,

3) the third nested factor is nested in a greater number ($k$) of factors than the first two,

4) the third nested factor is not nested in either of the first two nested factors,

is given by the expression:

$$
(3.24) \sum_{i=1}^{N-3} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=i+1}^{N-3} \sum_{k'=\max\{0,k-i-j\}}^{\min\{k-2j',2j\}} \sum_{k''=\max\{0,k-2j\}}^{k-k'-k''} \delta_{k',k''}
$$
The computed results for some values of \( N \) in expression (3.24) are given in Table 3.6.

The class of structures denoted 1.3 in Table 3.3 can be enumerated by the expression:

\[
\sum_{i=1}^{N-3} \min(i, N-i-3) \sum_{j'=0}^{N-2} \min(k, N-3-i-j') \sum_{k'=i+1}^{\max(0, k-i-j')} 1
\]

where \( \delta_{k'} = 1 \) if \( k' \leq \eta(S_{0}) = N-3-i-j' \),

and \( k-k' \leq \eta(S_{1}, N-1) = j' \),

\( = 0 \) otherwise.

**Lemma 3.7**

The number of structures possible with \( N \) factors having the properties:

1) three factors are not directly nested in the mean,
2) the first two such nested factors are nested in the same number \((i)\) of factors,
3) the third nested factor is nested in a greater number \((k\) where \(k = i\)) of factors than the first two nested factors,
4) the third nested factor is nested in one and only one of the other two nested factors,
is given by the expression:

\[
(3.26) \quad \sum_{i=1}^{N-3} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k'=i+1}^{N-2} \sum_{k''=\max(0,k-i-1-j')}^{\min(k-i-1-N-3-j')} \min(k-i-1-k',j')
\]

The computed results for some values of $N$ in expression (3.26) are given in Table 3.6.

The class of structures denoted 1.4 in Table 3.4 can be enumerated by the expression:

\[
(3.27) \quad \sum_{i=1}^{N-3} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k'=i+1}^{N-1} \delta_k
\]

where $\delta_k = 1$ if $k-i-j' - 2 \leq \gamma(S) = N-3-i-j'$

= 0 otherwise.

**Lemma 3.8**

The number of structures possible with $N$ factors having the properties

1) three factors are not directly nested in the mean,

2) the first two such nested factors are nested in the same number (i) of factors,

3) the third nested factor is nested in a greater number (k where $k = i$) of factors than the first two nested factors,
4) the third factor is nested in both of the other two nested factors,

is given by the expression:

\[
\sum_{i=1}^{N-3} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=i+1}^{N-1} \sum_{k'=\max(0,k-i-j'-2)}^{\min(k-i-j'-2,N-3-i-j')} 1
\]

(3.28)

The computed results for some values of \( N \) in expression (3.28) are given in Table 3.6.

By referring to Table 3.4, it can be seen that the class of structures containing two factors not directly nested in the mean, and characterized by \( i < j \), \( A_{N-2} \neq V_{N-1} \), is decomposable into six disjoint classes of structures by the inclusion of third factor not directly nested in the mean. For example, if \( N = 4, i = 1, \) and \( j = 2 \), the enumeration of the previous subsection yielded a partial structure of Class 2 given by:

![Class 2 Partial Structure](image)
If the structure is completed by the addition of a third factor, not directly nested in the mean, and the addition of the error term, the six resulting classes of structures 2.1 through 2.6 are illustrated by:

Class 2.1
1) \( i < j = k \),
2) \( A_{N-2} \not\in V_{N-1} \)
3) \( A_{N-2} \not\in V_N \)

none exists for \( N = 5 \)

Class 2.3
1) \( i < j < k \)
2) \( A_{N-2} \not\in V_{N-1} \)
3) \( A_{N-2} \not\in V_N \)
4) \( A_{N-1} \not\in V_N \)

Class 2.2
1) \( i < j = k \)
2) \( A_{N-2} \not\in V_{N-1} \)
3) \( A_{N-2} \in V_N \)

Class 2.4
1) \( i < j < k \)
2) \( A_{N-2} \not\in V_{N-1} \)
3) \( A_{N-2} \in V_N \)
4) \( A_{N-1} \not\in V_N \)
The six subclasses, as defined, are disjoint and exhaust possible completions of Class 2 partial structures. The first of these subclasses of structures, Class 2.1, is characterized by:

1) \( i \leq j = k \)
2) \( A_{N-2} \notin V_{N-1} \)
3) \( A_{N-2} \notin V_N \)
4) \( A_{N-1} \notin V_N \)

From the expressions in (3.20) and Table 3.4, the number of Class 2.1 structures possible, for a given value of \((N, i, j, j')\), can be expressed by

\[
\sum_{k'=0}^{j} \sum_{k''=0}^{j-k'} \sum_{k'''=0}^{j-k'-k''} \delta_{k,k',k''},
\]
where $\delta_{k',k'',k'''} = 1$ if $k' \leq \eta(S_0) = N-3-i-j'$,

$k'' \leq \eta(S_1,N-2) = \max\{0,i-j+j'^2\}$

$k''' \leq \eta(S_1,N-1) = j'$, and

$j-k'-k''-k''' \leq (S_2) = j-j'-\delta_{j'}^{j-1}$

= 0 otherwise.

However, some of the resulting structures will be isomorphic to those resulting from the completion of a partial structure with parameters $(N, i, j^*, j'^*)$. For example, consider the two partial structures:

$N = 6, i = 1, j = 3, j' = 2$ \hspace{1cm} $N = 6, i = 1, j = 3, j' = 3$

![Diagram](image)

The structure on the left can be completed using parameters $k' = 1$, $k'' = 0$, $k''' = 2$. Then, this yields the complete seven factor structure:
The second partial structure can be completed using parameters \( k' = 0 \), \( k'' = 1 \), \( k''' = 2 \). This yields the complete seven factor structure

\[ N = 7, i = 1, j = 3, j' = 2, k' = 1, k'' = 0, k''' = 2 \]

The second partial structure can be completed using parameters \( k' = 0 \), \( k'' = 1 \), \( k''' = 2 \). This yields the complete seven factor structure

\[ N = 7, i = 1, j = 2, j' = 3, k' = 0, k'' = 1, k''' = 2 \]

However, the two resulting structures are isomorphic. This difficulty is avoided by modifying (3.29) and eliminating the enumeration of structures such that \( k' + k''' > j \). The structure given in (3.30) becomes inadmissible with the preceding restriction.

In summary, the result is given by:

**Lemma 3.9**

The number of non-isomorphic structures containing \( N \)-factors, having the properties
1) three factors are directly nested in factor other than the mean,

2) if i, j and k, are the number of nestors (excluding the mean) of the three factors not directly nested in the mean, then
   \[ i < j = k, \]

3) \( A_{N-2} \neq V_{N-1} \)

4) \( A_{N-2} \neq V_N \)

is given by

\[
(3.32) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \min(i, N-i-3) \sum_{j'=0}^{\min(j, N-3-i-j')} \sum_{k'=0}^{\min(j, j')} \sum_{l'=0}^{\min(l, j')} \sum_{k''=0}^{\min(k'', j'')} \sum_{k'''=0}^{\min(k''', j''')} \delta_{j', k', k'', k'''}
\]

where \( \delta_{j', k', k'', k'''} = 1 \) if \( j'-k''-k'''' \leq \gamma(S_2) 2 j'-j''-\delta_{j', i} \),

and \( k' - j' \leq k'''' \),

\[ = 0 \] otherwise.

The computed results for some values of N in the expression (3.32) are given in Table 3.6.

For the class of structure denoted 2.2 in Table 3.4, the number of structures as a function of N is given by the expression:
The number of structures with $N$ factors, having the properties

1) three factors are not directly nested in the mean,

2) if $i$, $j$ and $k$, are the numbers of nestors (excluding the mean) of the three factors not directly nested in the mean, then $i < j = k$,

3) $A_{N-2} \not\in V_{N-1}$

4) $A_{N-2} \in V_{N}$,

is given by

\[
(3.34) \sum_{i=1}^{N-3} \sum_{j=1+i}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k'=0}^{\min(j-1,N-3-i-j')} \delta_{i,j,j',k'}
\]

where $\delta_{i,j,j',k'} = 0$ if $k' \leq \eta(S_0) = N-3-i-j'$, and $j-1-k' \leq \eta(S_{N-1}) = j'$, and $\delta_{i,j,j',k'} = 0$ otherwise.

Lemma 3.10

The computed results for some values of $N$ from expression (3.34) are given in Table 3.6.
For the class of structures denoted by 2.3 in Table 3.4, the number of structures as a function of \( N \) is given by the expression

\[
\sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \min(i,N-i-3) \sum_{k=j+1}^{N-1} \sum_{k'=0}^{k-k'} \sum_{k''=0}^{k-k''} \delta_{k',k'',k'''} = 0
\]

where \( \delta_{k',k'',k'''} = 1 \) if

\[
k' \leq \gamma(S_0) = N-3-i-j',
\]

\[
k'' \leq \gamma(S_{1,N-2}) = i - j + j',
\]

\[
k''' \leq \gamma(S_{1,N-1}) = j',
\]

\[
k-k'-k''-k''' = \gamma(S_2) = j-3,
\]

\( = 0 \) otherwise.

**Lemma 3.11**

The number of structures with \( N \) factors having the properties

1) three factors are not directly nested in the mean,

2) if \( i, j \) and \( k \), are the number of nestors (excluding the mean) of the three factors not directly nested in the mean, then \( i < j < k \),

3) \( A_{N-2} \neq V_{N-1} \),

4) \( A_{N-2} \neq V_N \),

5) \( A_{N-1} \neq V_N \),

is given by

\[
\sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \min(i,N-i-3) \sum_{k=j+1}^{N-1} \min(k,N-3,i-j') \sum_{k'=0}^{\max(0,k-i-j')} \delta_{k',k'',k'''} = 0
\]

...
\[
\sum_{k''=\max(0, k-k'-j)} \min(k-k', i-j+j') \quad \sum_{k'''=\max(0, k-k'-k''-j+j')} \min(k-k''-k''', j+j')
\]
\[
\sum_{k''''=\max(0, k-k''-k'''-j+j')} \min(k-k''-k''', j-j')
\]

The computer results for some values of \(N\) obtained from expression (3.36) are given in Table 3.6.

For the class of structures denoted by 2.4 in Table 3.4, the number of structures as a function of \(N\) is given by the expression:

\[(3.37) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k'=0}^{\min(k-i-1, N-1)} \delta_{k,k'}
\]

where \(\delta_{k,k'} = 1\) if \(k' \leq \eta(S_0) = N-3-i-j'\), and \(k'i-1-k' \leq j'\),

\(= 0\) otherwise.

**Lemma 3.12**

The number of structures with \(N\) factors having the properties

1) three factors are not directly nested in the mean,

2) if \(i, j\) and \(k\), are the number of nestors (excluding the mean) of the three factors not directly nested in the mean, then \(i < j < k\),

3) \(A_{N-2} \notin V_{N-1}'\),

4) \(A_{N-2} \subseteq V_{N}'\),

...
is given by

\[
N-3 \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-2} \min(i, N-i-3) \sum_{j'=0}^{N-1} \sum_{k=j+1}^{N-1} \min(k-i-1, N-3-i-j') \sum_{k'=\max(0, k-i-1-j')}^{k'} \min(k-i-1-k', j') \sum_{k''=\max(0, k+2-N+j')}^{k''} 
\]

The computer results for some values of \( N \) obtained from expression (3.38) are given in Table 3.6.

For the class of structures denoted 2.5 in Table 3.4, the number of structures as a function of \( N \) is given by:

\[
N-3 \sum_{i=1}^{N-2} \sum_{j=i+1}^{N-2} \min(i, N-i-3) \sum_{j'=0}^{N-1} \sum_{k=j+1}^{N-1} \sum_{k=0}^{\min(k-i-1,k'-j)} \min(k-i-1-k', j') \sum_{k''=\max(0, k+2-N+j')}^{k''} 
\]

where \( \delta_{k,k'} = 0 \) if \( k' \leq \gamma(S_0) = N-3-i-j' \),

and \( k-k'-j-l \leq \gamma(S_1, N-2) = \max(0, i-j+j') \)

\( = 0 \) otherwise.

Lemma 3.13

The number of structures with \( N \) factors having the properties

1) three factors are not directly nested in the mean,

2) if \( i, j \) and \( k \), are the number of nestors (excluding the mean) of the three factors not directly nested in the mean, then
i \leq j \leq k,

3) A_{N-2} \notin V_{N-1},

4) A_{N-2} \notin V_N,

5) A_{N-1} \in V_N.

is given by

\begin{equation}
(3.40) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=j+l}^{N-1} \sum_{k'=\max(0,k-j-1-\max(0,i+j'))}^{\min(k-j-1-k',0)}
\end{equation}

The computer results for some values of N obtained from (3.40) are given in Table 3.6.

For the class of structures denoted 2.6 in Table 3.4 the number of structures as a function of N is given by

\begin{equation}
(3.41) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=i+j'+1}^{N-1} \sum_{k'=0}^{\min(k-i-j'-2,0)} \delta_{k,k'}
\end{equation}

where \( \delta_{k,k'} = 1 \) if \( k' \leq \chi(S_0) = N-3-i-j' \),

\( = 0 \) otherwise.

Lemma 3.14

The number of structures with N factors, having the properties

1) three factors are not directly nested in the mean,

2) if i, j and k, are the number of nestors (excluding the mean)
of the three factors not directly nested in the mean; then

\[ i < j < k, \]

3) \( A_{N-2} \notin V_{N-1} \),

4) \( A_{N-2} \notin V_N \),

5) \( A_{N-1} \notin V_N \)

is given by

\[
(3.42) \quad \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} (N-1-j'-2).
\]

The computer results for some values of \( N \) obtained from (3.42) are given in Table 3.6.

For the class of structures denoted 3.1 in Table 3.5, the number of structures as a function of \( N \) is given by:

\[
(3.43) \quad \sum_{i=1}^{N-3} \sum_{k=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k'=0}^{j'-1} \delta_{k'},
\]

where \( \delta_{k'} = 1 \) if \( k' \leq \gamma(S_0) = N-3-i-j' \), and \( j'-k'-i-1 \leq \gamma(S_1,N-1) = j' \),

= 0 otherwise.

**Lemma 3.15**

The number of structures with \( N \) factors, having the properties

1) three factors are not directly nested in the mean,

2) if \( i, j \) and \( k \), are the number of nestors (excluding the mean)
of the three factors not directly nested in the mean, then
\[ i < j < k, \]
3) \( A_{N-2} \in V_{N-1} \),
4) \( A_{N-2} \in V_N \),

is given by

\[
(3.44) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k''=0}^{\min(j',N-3-i-j')} 1.
\]

The computer results for some values of \( N \), obtained from (3.44) are given in Table 3.6.

For the class of structures denoted 3.2 in Table 3.5, the number of structures as a function of \( N \) is given by

\[
(3.45) \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=0}^{N-1} \sum_{k'=0}^{k} \sum_{k''=0}^{k-k'} \delta_{k',k''},
\]

where \( \delta_{k',k''} = 1 \) if \( k' \leq \gamma(S_0) = N-3-i-j' \),

and \( k'' \leq \gamma(S_{1,N-1}) = j' \),

and \( k-k'-k'' \leq \gamma(S_1) = i \)

= 0 otherwise.

**Lemma 3.16**

The number of structures with \( N \) factors, having the properties

1) three factors are not directly nested in the mean,
2) if \( i, j \) and \( k \), are the number of nestors (excluding the mean)
of the three factors not directly nested in the mean, then

\[ i < j < k, \]

3) \( A_{N-2} \not\subseteq \mathcal{V}_{N-1} \)

4) \( A_{N-2} \not\subseteq \mathcal{V}_N \)

5) \( A_{N-1} \not\subseteq \mathcal{V}_N \)

is given by

\[
(3.46) \quad \sum_{i=1}^{N-3} \sum_{k=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=j+1}^{N-1} \sum_{k'=\max(0,k-j-i)}^{\min(k,N-3-i-j')} \sum_{k''=\max(0,k-N+3+j')}^{\min(k-k',j')} \sum_{k'''=\max(0,k-N+3+j')}^{1} 1.
\]

The computer results for some values of \( N \), obtained from (3.46) are given in Table 3.6.

For the class of structures denoted 3.3 in Table 3.5, the number of structures as a function of \( N \) is given by

\[
(3.47) \quad \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \sum_{k=j+1}^{N-1} \sum_{k'=0}^{k-i+1} \delta_{k''},
\]

where \( \delta_{k''} = 0 \) if \( k' \leq \eta(S_0) = N-3-i-j' \), and \( k-i-l-k' \leq \eta(S_1, N-1) = j' \),

\[ = 0 \] otherwise.

Lemma 3.17

The number of structures with \( N \) factors, having the characteristics
1) three factors are not directly nested in the mean,
2) if $i$, $j$ and $k$, are the number of nestors (excluding the mean) of the three factors not directly nested in the mean, then $i < j < k$,
3) $A_{N-2} \notin V_{N-1}$,
4) $A_{N-2} \notin V_N$,
5) $A_{N-1} \notin V_N$,

is given by

$$
(3.48) \quad \sum_{i=1}^{N-3} \sum_{j=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \min(k-i-1,N-3-i-j') \sum_{k'=j+1}^{\min(k-i-1,N-1-j')} \sum_{k''=\max(0,k+i-1-j')}^{\max(0,k+2-N+j')} \min(k-i-i-k',j').
$$

The computer results from evaluation of expression (3.48) for some values of $N$ are given in Table 3.6.

**Lemma 3.17**

For the class of structures denoted 3.4 in Table 3.5, the number of structures as a function of $N$ is given by

$$
(3.49) \quad \sum_{i=1}^{N-3} \sum_{k=i+1}^{N-2} \sum_{j'=0}^{\min(i,N-i-3)} \min(i,N-i-3) \sum_{j'=0}^{(N-j-1)}.
$$

The computer results for some values of $N$ in expression (3.49) are given in Table 3.6.
It is clear from the expressions in (3.22), (3.24), (3.26), (3.28), (3.32), (3.34), (3.36), (3.38), (3.40), (3.42), (3.44), (3.46), (3.48), and (3.49) that each structure containing N factors with three factors not directly nested in $\mu$ has a characterization in terms of the tuple $(i, j, j', k, k', k'', k''', k''''')$ suitably restricted.
Table 3.6. Number of structures containing three factors not directly nested in the mean, by classes of structures as defined in Table 3.3, 3.4 and 3.5

<table>
<thead>
<tr>
<th>Structure Class</th>
<th>Number of Factors</th>
<th>1.1</th>
<th>1.2</th>
<th>1.3</th>
<th>1.4</th>
<th>2.1</th>
<th>2.2</th>
<th>2.3</th>
<th>2.4</th>
<th>2.5</th>
<th>2.6</th>
<th>3.1</th>
<th>3.2</th>
<th>3.3</th>
<th>3.4</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td></td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>27</td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>7</td>
<td>8</td>
<td>12</td>
<td>9</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>6</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>10</td>
<td>88</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>13</td>
<td>22</td>
<td>25</td>
<td>17</td>
<td>18</td>
<td>10</td>
<td>12</td>
<td>23</td>
<td>18</td>
<td>14</td>
<td>13</td>
<td>5</td>
<td>11</td>
<td>20</td>
<td>221</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td>22</td>
<td>48</td>
<td>44</td>
<td>28</td>
<td>43</td>
<td>21</td>
<td>43</td>
<td>58</td>
<td>43</td>
<td>31</td>
<td>22</td>
<td>15</td>
<td>24</td>
<td>35</td>
<td>477</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>35</td>
<td>93</td>
<td>73</td>
<td>43</td>
<td>88</td>
<td>38</td>
<td>117</td>
<td>123</td>
<td>87</td>
<td>59</td>
<td>34</td>
<td>36</td>
<td>46</td>
<td>56</td>
<td>928</td>
</tr>
<tr>
<td>10</td>
<td></td>
<td>52</td>
<td>165</td>
<td>112</td>
<td>62</td>
<td>164</td>
<td>64</td>
<td>273</td>
<td>234</td>
<td>160</td>
<td>102</td>
<td>50</td>
<td>74</td>
<td>80</td>
<td>84</td>
<td>1676</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td>75</td>
<td>275</td>
<td>166</td>
<td>86</td>
<td>285</td>
<td>100</td>
<td>569</td>
<td>410</td>
<td>273</td>
<td>164</td>
<td>70</td>
<td>138</td>
<td>130</td>
<td>120</td>
<td>2860</td>
</tr>
<tr>
<td>12</td>
<td></td>
<td>105</td>
<td>435</td>
<td>235</td>
<td>115</td>
<td>469</td>
<td>150</td>
<td>1096</td>
<td>676</td>
<td>439</td>
<td>250</td>
<td>95</td>
<td>238</td>
<td>200</td>
<td>165</td>
<td>4667</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>143</td>
<td>662</td>
<td>325</td>
<td>150</td>
<td>740</td>
<td>215</td>
<td>1979</td>
<td>1061</td>
<td>673</td>
<td>365</td>
<td>125</td>
<td>380</td>
<td>295</td>
<td>220</td>
<td>7341</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td>190</td>
<td>973</td>
<td>436</td>
<td>191</td>
<td>1126</td>
<td>300</td>
<td>3400</td>
<td>1601</td>
<td>998</td>
<td>515</td>
<td>161</td>
<td>603</td>
<td>420</td>
<td>286</td>
<td>11200</td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>248</td>
<td>1391</td>
<td>575</td>
<td>239</td>
<td>1664</td>
<td>406</td>
<td>5597</td>
<td>2337</td>
<td>1434</td>
<td>706</td>
<td>203</td>
<td>903</td>
<td>581</td>
<td>364</td>
<td>16648</td>
</tr>
</tbody>
</table>
3. Extensions of the induction procedure

Part of the results of parts 1 and 2 of this section yield the following numbers:

<table>
<thead>
<tr>
<th>N</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Total Number of Structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>7</td>
<td>5</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>16</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4</td>
<td>15</td>
<td>27</td>
<td>16x</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>63</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>5</td>
<td>27</td>
<td>88</td>
<td>x</td>
<td>63x</td>
<td>-</td>
<td>-</td>
<td>319x</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>6</td>
<td>43</td>
<td>221</td>
<td>x</td>
<td>x</td>
<td>319x</td>
<td>-</td>
<td>x</td>
</tr>
<tr>
<td>8</td>
<td>1</td>
<td>7</td>
<td>64</td>
<td>477</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

The preceding sections carried the induction only through \( q = 3 \). However, the numbers marked \( x \) above, can be obtained from the following.

**Lemma 3.18**

The number of structures containing \( N \) factors and having \( N-1 \) factors nested in factors other than the mean, is equal to the total number of structures containing \( N-1 \) factors.

Clearly if only one factor is nested directly in the mean, then the possible structures are of the form:

\[
\begin{array}{c}
\nu \\
A_1
\end{array}
\]

(some arrangement of \( N-1 \) nodes).
The number of such possible arrangements is the number of structures containing $N-1$ factors.

The number 319 marked with $\approx$ in the preceding table, is due to Gilbert. In order, to obtain the number of structures containing six factors it is necessary to extend the induction process to the case when $q = 4$.

Consider any partial structure, containing three factors not directly nested in the mean. It is possible to generalize the partition of nodes into sets of like nodes, by defining:

- $S_0$ = the set of nodes, directly nested in the mean and not attached diagrammatically to $A_{N-3}$, $A_{N-2}$, and $A_{N-1}$.
- $S_{1,N-3}$ = the set of nodes, directly nested in the mean, and attached to $A_{N-3}$, but not to $A_{N-2}$ and $A_{N-1}$.
- $S_{1,N-2}$ = the set of nodes, directly nested in the mean, and attached to $A_{N-2}$, but not to $A_{N-3}$ and $A_{N-1}$.
- $S_{1,N-1}$ = the set of nodes, directly nested in the mean, and attached to $A_{N-1}$, but not to $A_{N-3}$ and $A_{N-1}$.
- $S_{2,N-3,N-2}$ = the set of nodes, directly nested in the mean, and attached diagrammatically to $A_{N-3}$ and $A_{N-2}$, but not to $A_{N-1}$.
- $S_{2,N-3,N-1}$ = the set of nodes, directly nested in the mean, and attached diagrammatically to $A_{N-3}$ and $A_{N-1}$, but not to $A_{N-2}$.
- $S_{2,N-2,N-1}$ = the set of nodes directly nested in the mean, and attached diagrammatically to $A_{N-2}$ and $A_{N-1}$, but not to $A_{N-3}$. 
\[ S_{3,N-3,N-2,N-1} \] = the set of nodes directly nested in the mean, and
attached diagrammatically to \( A_{N-3} \), \( A_{N-2} \), \( A_{N-1} \).

\[ S_{4,N-3} \] = the set containing \( A_{N-3} \).
\[ S_{4,N-2} \] = the set containing \( A_{N-2} \).
\[ S_{4,N-1} \] = the set containing \( A_{N-1} \).

Therefore \( q = 4 \) requires a partition of the \( N-1 \) nodes into 11 sets. In general, an arbitrary value of \( q \) would require:

\[
\sum_{i=0}^{q-1} \binom{q-1}{i} + (q-1) = 2^{q-1} + q - 1,
\]
sets of like nodes.

For the case \( q = 4 \), it is possible to determine the number of nodes contained by each of the preceding sets, as a function of \( N, i, j, j', k, k', k'', k''' \), parameters which characterize each partial structure. It is tedious, but not difficult to divide each class of structures obtained, and extend the equations derived in part 2 of this section by additional summations. The possibility that the computer can be programmed to extend the results for arbitrary \( q \) is currently being investigated.
IV. IDENTIFICATION OF ESTIMABLE FUNCTIONS IN ARBITRARY CLASSIFICATION ARRANGEMENTS AND NON-ORTHOGONAL ANALYSIS OF VARIANCE

A. The General Problem

The simplicity of the logical computational algorithm described in Chapter II is possible only when the data possess sufficient structure to be classified as balanced and complete, in accordance with Definitions 2.10 and 2.9. Data resulting from designs, such as latin squares, lattices, and balanced or partially balanced incomplete designs, etc., fall into the category of balanced and incomplete with respect to the same definitions. Although this work does not contain specific computational algorithms for the analysis of balanced incomplete data, Chapter II reviewed and commented on methods for computations suggested by the work of Zyskind (36), White (26), and Nelder (21). Another category of experimental design data is that which may be termed unbalanced complete data. For this class, various computational algorithms have been implemented using general linear multiple regression techniques on a full-rank reparametrization of the model. Examples of these are Bock (1), Dixon (7), Federer and Zelen (9), Powlkes (10), Wilkinson (32). Additionally, the AARDVARK system discussed in Chapter II also contains an algorithm for non-orthogonal analysis of variance through multiple linear regression techniques. A description of the AARDVARK non-orthogonal algorithm, and a critical evaluation of its method of analysis when one or more cells are empty, are contained in Section B.
The objective sought by the designers of these computer programs is to provide efficiently the classical interpretations of the data via estimation, hypothesis tests, and confidence intervals. Each of the categories of data mentioned above, namely, balanced complete, balanced incomplete data resulting from a standard design, and unbalanced complete data, has one basic and extremely important characteristic in common. This is, that there is a priori knowledge of side conditions and restrictions which lead to a basis of interest for the space of all independently linearly estimable parametric functions. This knowledge of what is estimable and what is not makes possible the algorithms which were referenced in the preceding.

It is frequently the case that the data available are unbalanced and also incomplete. In arbitrary incomplete fixed factorial arrangements, there is generally no a priori knowledge of functions estimable with the data. Indeed, when the model fitted contains interaction parameters the functions that may be estimated, and therefore tested, about main effects are inextricably related to the restrictions one imposes. If the models to be fitted are additive, the interpretation of the data would logically seem to require an answer to two questions. First, what functions are independently estimable within the parameters of each factor of classification. Second, if the first set does not span a basis for all estimable functions, then what additional independent functions can be estimated with the available data. The latter set of independent estimable functions will contain parameters from
more than one factor of classification. Thus, the experimenter will have to decide if they contain any information of use. The attainment of a basis as described above is a problem of major concern in this and the next chapter. This chapter exhibits the problems that arise and surveys critically the limited literature available on the analysis of variance of arbitrary incomplete classification data. The solution proposed is designed to be suitable for implementation on a digital computer.

It is supposed that there exists a set of observations \( \{ y_{ij} \}_{i=1}^{n} \) such that each can be classified as belonging to a class or cell. Each cell contains an attribute or attributes somehow measured by the observations in that cell subject to a random error independent for each observation, having zero expectation and constant variance. Let \( \{ \mu_i \}_{i=1}^{P} \) denote the set of cell attributes. It is possible to arrange the observations and to represent the situation by the model:

\[
y_{ij} = \mu_i + e_{ij}, \quad i = 1, 2, \ldots, p; \quad j = 0, 1, \ldots, n_i
\]

where the \( n_i \)'s are the numbers of observations in the \( i^{th} \) cell. A necessary and sufficient condition for estimability of any \( \mu_i \) is that \( n_i \) be greater than zero. The best linear unbiased estimator of any estimable \( \mu_i \) is then given by the corresponding cell mean.

In planned factorial experiments the cells may be defined by the levels of controlled factors. In the analysis of data which are not the direct result of planned experimentation, it is often convenient to
superimpose a similar type of factorial structural arrangement, because it is desired to interpret the observed variation in terms of the levels of factors hypothesized to contribute to the observed variability. This results in a decomposition of each cell response into effects and interaction parameters. In the notation of Federer and Zelen (9), an \( n \)-way crossed experimental arrangement which admits all possible interactions results in a decomposition of each \( \mu_i \) given by:

\[
\mu_i = \mu + \sum_{s=1}^{n} a_s(j_s) + \sum_{r=1}^{n-1} \sum_{s=r+1}^{n} a_{rs}(j_r, j_s) + \\
+ \sum_{q=1}^{n-2} \sum_{r=t+1}^{n-1} \sum_{s=t+1}^{n} a_{qrs}(j_q, j_r, j_s) + \ldots + a_{123\ldots n}.
\]

With the model 4.1, the functions estimable are given simply by the rule: only functions of \( \mu_i \)'s corresponding to \( n_i \)'s > 0 are estimable, and the best linear unbiased estimate of each is same linear function of the corresponding cell means. If the \( \mu_i \) are decomposed as in 4.2, what is estimable is not obvious.

An initial complication introduced by a decomposition of cell attributes into classificatory parameters is exhibited by the obvious fact that with the additive decomposition:

\[
\mu_i = \mu + \sum_{s=1}^{n} a_s(j_s)
\]

it may be possible to estimate a \( \mu_i \) corresponding to an unobserved cell. This is never true if the decomposition involves any interaction.
parameters.

Suppose the model is as given in (4.1) and (4.2), and the data available have arbitrary incidence. In the following, it is assumed that the experimenter is interested in:

1) determining the independent parametric functions in the main effects that are estimable with the data,

2) determining the independent parametric functions in the interaction parameters that are estimable with the data,

3) determining the independent functions in the main effects, and interactions that are not estimable with the available data, but would have been estimable had all cells been observed.

It is further assumed that the experimenter is willing to impose only those parametric restraints which arise naturally in the development of the derived linear model. Hence, with respect to equation (4.2), it may be assumed that:

1) for every \( s \), \( \sum_{j_s} a_s(j_s) = 0 \);

(4.3) 2) for every \( (r,s) \), \( \sum_{j_r} \sum_{j_s} a_{rs}(j_r, j_s) = 0 \);

and \( \sum_{j_s} a_{rs}(j_r, j_s) = 0 \); etc.
B. A General Method of Computing the Non-orthogonal Analysis of Variance

In matrix notation, the models under consideration are of the form:

\( y = X\beta + Z\delta + e, \)

where \( y \) is an \( n \times 1 \) observational vector, \( X \) is an \( n \times p \) matrix of \( 0's \), and 1's corresponding to the coefficients of the classification part of the model, \( \beta \) is a \( p \times 1 \) vector of classification parameters, \( Z \) is an \( n \times q \) matrix of coefficients corresponding to the \( q \times 1 \) vector of covariate parameters in \( \delta \). The unknown parameters in \( \beta \) and \( \delta \) are assumed fixed, and the \( n \times 1 \) vector \( e \) is assumed random having multivariate distribution, such that \( E(e) = 0 \) and \( E(e'e') = \sigma^2 I \). The method here described is applicable only when \( \bar{\beta}(X) \cap \bar{\beta}(Z) = \phi \), where \( \phi \) denotes the column vector space and is the empty set. The problems that result when this assumption of disjoint spaces does not hold are numerical rather than statistical. The objective in the present section is the efficient reparametrization of 4.4 to a model:

\( y = W\Theta + Z\delta + e \)

where \( \Theta \) is now a vector of \( r \) fixed unknown parameters, and \( r \) is the rank of \( X \). Such a full rank reparametrization always exists and for arrangements of maximal rank may be obtained computationally by choosing a known matrix \( T \) such that
(4.6) $\beta = T \Theta$ \quad and \quad $W = XT$.

It is well known that if $\vec{\beta}(X) = \vec{\beta}(W)$, there is a one-to-one correspondence between estimable linear functions with the model (4.4) and estimable linear functions with the model in (4.5). The following provides a computational algorithm useful in obtaining such a full rank reparametrization. Further, it will be seen that it is unnecessary to construct the matrices of the model (4.5), but rather that it is possible to obtain directly the corresponding normal equations.

(4.7) $W'W\Theta + W'Z\delta = W'y$

$Z'W\Theta + Z'Z\delta = Z'y$.

The computer program requires the following minimal inputs:
(a) the classification structure of the data,
(b) the maximum subscript range of each index associated with a factor,
(c) the level of each factor associated with each observation,
(d) the corresponding observational vectors ($y$ and $z$).

It is very important that the $X$ matrix never be required. In general, this $n \times p$ matrix may be significantly larger than the resulting matrix $W'W$. Since $W'W$ is sufficient for the analysis of variance, the prescribed procedure computes $W'W$ without forming $X$ or $W$.

In the remainder of this section, the model is taken to be one without covariates, i.e. with $Z = 0$. Further it is supposed that the
program has available the full structure matrix, as defined in Section E of Chapter II, and a vector R containing the maximum subscript range of each index. The procedure described is essentially that followed by the AARDVARK system for non-orthogonal data with slight modifications.

Let the vector P contain a code identifying each term in the model, such that

\[ P(i) = -1 \] for the mean and other factors directly nested in the mean,
\[ = 0 \] for factors directly nested in factors other than the mean,
\[ = 1 \] for interaction model terms.

Next, it is desirable to identify the columns of the matrix X as belonging to a certain model term. For this purpose, we construct two arrays LL and LH containing the lowest and highest column numbers belonging to each model term. A matrix, LN, is constructed containing the position of each subscript associated with each model term, and an array, N, is formed such that it contains the number of non-zero entries in the corresponding row of the matrix.

The preceding sequence is illustrated in the following example.

**COMPUTER SPECIFICATIONS**

\[
\begin{align*}
\text{FACTOR (1)} &= 'A', & \text{LIMITS (1)} &= 3, \\
\text{FACTOR (2)} &= 'B', & \text{LIMITS (2)} &= 4, \\
\text{FACTOR (3)} &= 'AB(C)', & \text{LIMITS (3)} &= 2, \\
\text{FACTOR (4)} &= 'B(D)', & \text{LIMITS (4)} &= 2,
\end{align*}
\]
FACTOR (5) = 'ABCD(E)'
LIMITS (5) = 6.

Note that the structure specified is that given in Section E of Chapter II. The mean is not specified but is assumed. The LIMITS specification gives the maximum subscript range of each index associated with each factor.

COMPUTED CONTROL ARRAYS

ITALICS STRUCTURE

<table>
<thead>
<tr>
<th>MATRIX</th>
<th>F</th>
<th>II</th>
<th>IH</th>
<th>N</th>
<th>IN</th>
</tr>
</thead>
<tbody>
<tr>
<td>[ \begin{bmatrix} \mu \ A \ B \ AB \ AB(G) \ B(D) \ B(AD) \ AB(CD) \ ABCD(E) \end{bmatrix} ]</td>
<td>1 1 1 1 1</td>
<td>-1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0 1 1 1 1</td>
<td>-1</td>
<td>2</td>
<td>4</td>
<td>1</td>
<td>1 0 0 0</td>
</tr>
<tr>
<td>1 0 1 1 1</td>
<td>-1</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>2 0 0 0</td>
</tr>
<tr>
<td>0 0 1 1 1</td>
<td>1</td>
<td>9</td>
<td>20</td>
<td>2</td>
<td>1 2 0 0</td>
</tr>
<tr>
<td>0 0 0 1 1</td>
<td>0</td>
<td>21</td>
<td>44</td>
<td>3</td>
<td>1 2 3 0</td>
</tr>
<tr>
<td>1 0 1 0 1</td>
<td>1</td>
<td>45</td>
<td>52</td>
<td>2</td>
<td>2 4 0 0</td>
</tr>
<tr>
<td>0 0 1 0 1</td>
<td>1</td>
<td>53</td>
<td>84</td>
<td>3</td>
<td>1 2 4 0</td>
</tr>
<tr>
<td>0 0 0 0 1</td>
<td>1</td>
<td>85</td>
<td>276</td>
<td>4</td>
<td>1 2 3 4</td>
</tr>
<tr>
<td>0 0 0 0 0</td>
<td>-</td>
<td>--</td>
<td>-</td>
<td>-</td>
<td>- - - -</td>
</tr>
</tbody>
</table>

With the preceding control vectors, it is possible to examine each observation and construct the corresponding row of the X matrix. The observation record is taken of the form exemplified in the following.

(4.16) INDICES

\[ \begin{array}{cccc}
I & J & K & L & M \\
1 & 1 & 1 & 1 & 1
\end{array} \]

DEPENDENT VARIABLES

\[ \begin{array}{cccc}
Y_1 & Y_2 & \ldots & Y_k \\
2.8 & 3.2 & \ldots & 10.5
\end{array} \]

INDEPENDENT VARIABLES

\[ \begin{array}{cccc}
Z_1 & Z_2 & \ldots & Z_q \\
3.2 & 1.0 & \ldots & 8.5
\end{array} \]

The row of the X matrix is constructed by examining the control
arrays, one model term at a time, adding 1's in the appropriate locations of a vector initialized to contain all zeroes. The location of the 1 for model term $J$ is computed as

$$\frac{N(J)}{J} \sum_{I=1}^{I} \text{INDEX}(IN(I)) + LL(J).$$

(4.17)

Computation of all the partial means for each of the variables requires a vector of the same dimension as the rows of $X$ for each variable. Then, for each observational record, the value of the variables are accumulated into the positions, computed with (4.17), of that vector corresponding to the variable. The row vectors of $X$ are accumulated for each observation into a vector $S_X$. The entries in $S_X$ are then the appropriate divisors for computation of the means.

Note that the above procedure requires only one row of $X$ in core at a time. Total core required for computation of the means is $(n + q + p)2$, where $n$ is the number of dependent variables, $q$ is the number of independent variables, and $p$ is the dimension of the vector $\beta$.

In the present discussion, we assume that the divisors contained in the vector $S_X$ are all positive, i.e., there are no empty cells. The observation data are again examined and for each observation the corresponding row of the full rank matrix $W$ is constructed. This requires, first, that the control vectors $L$ and $LH$ be modified to correspond to the column space of $W$. If the modified vectors are called $LW$ and $LHW$, respectively, then for the preceding example these would be as follows.
The rules for forming the row of $W$ corresponding to a given observation are given in the following. The rules are given for each type of model term and consist of combining the indices for the observation in question with the control vectors whose construction was previously described.

Let $\text{INDEX}$ be the vector containing the indices associated with a particular observation. With respect to a given model term, let $I$ be the indices associated with the rightmost bracket, $J$ be the set of indices associated with the non-rightmost bracket symbols and let $\eta(I)$, $\eta(J)$ be, respectively, the number of indices. The rules for forming a row of $W$ are as follows.

1) For the mean.

Enter a 1 in position 1.
2) For factors directly nested in the mean.

Enter a 1 in position

\[ \text{INDEX}(I) + \text{LLW} \text{ if } \text{INDEX}(I) = \text{LIMIT}(I). \]

Enter a -1 in the positions

\[ \text{LLW through LLW} \text{ if } \text{INDEX}(I) = \text{LIMIT}(I). \]

3) For factors directly nested in factors other than the mean:

Enter a 1 in the position

\[
\sum_{j=1}^{(j-1)} \left[ \text{INDEX}(J) - 1 \frac{\eta(j)}{\sum_{k=J+1}^{\text{LIMIT}(K)}} + \text{INDEX}(\eta(j)) + \text{INDEX}(I) +
\right.

+ \text{LLW} - 1,

if \text{INDEX}(I) < \text{LIMIT}(I).

Enter a -1 in the positions

\[
\sum_{j=1}^{(j=1)} \left[ \text{INDEX}(J) - 1 \frac{\eta(j)}{\sum_{k=J+1}^{\text{LIMIT}(K)}} + \text{INDEX}(J) + \text{LLW}
\right.

through

\[
\sum_{j=1}^{(j=1)} \left[ \text{INDEX}(J) - 1 \frac{\eta(j)}{\sum_{k=J+1}^{\text{LIMIT}(K)}} + \text{INDEX}(J) + \text{LIMITS}(I) +
\right.

+ \text{LLW} - 1,

if \text{INDEX}(I) = \text{LIMITS}(I).

4) For interactions:

Enter in
\[
\frac{1}{\prod_{i} (\text{LIMIT}(i) - 1)} \left[ \sum_{j} \left( \frac{\eta(j)}{\text{LIMIT}(i)} \right) \left( \text{INDEX}(j) - 1 \right) \right] + \text{INDEX}(\eta(j)) + \sum_{i} \left( \frac{\eta(i)}{\text{LIMIT}(i) - 1} \right) \left( \text{LIMIT}(i) - 1 \right)
\]

the value

\[
\prod_{i} \delta_{\text{LIMITS}(i), \text{INDEX}(i)}
\]

where

\[
\delta_{\text{INDEX}(i)} = 1 \quad \text{if} \quad \text{INDEX}(i) \lt \text{LIMITS}(i)
\]

\[
= -1 \quad \text{if} \quad \text{INDEX}(i) = \text{LIMITS}(i).
\]

By applying the preceding rules we complete the construction of the r dimensional row vector of W that corresponds to a particular observation. Since \(W'W\) is symmetric, significant reduction of the core utilization can be achieved by storing only the upper triangular part of the matrix. This requires a one dimensional array to store matrix elements and a vector of pointers containing the position of the first element for each row. The array in which the elements of \(W'W\) are stored, here called V, requires \(r(r + 1)/2\) positions. The position in the vector V of the first element of the \(m\)-th row of \(W'W\) is given by the value of the expression:
\[ IPOS(m) = 1 + \sum_{i=1}^{m} (r - i + 1). \]

If \( w_i \) is the \( i \)-th element of the row vector of \( W \) for a given observation, then the procedure for forming \( W'W \) in \( V \) consists of computing the sums of squares and cross products of the \( w_i \) successively. Then, the products \( w_i w_i' \) are computed for all values \( i \leq i' \),

and \( V(IPOS(m) + i') = V(IPOS(m) + i) + w_i w_i' \).

Note that the above yields \( W'W \) compactly stored and requires only one row of \( W \) in core at any given time. Hence, the size of the problem is restricted only by the rank of the design matrix and not by the number of observations.

It is necessary to invert \( W'W \) to obtain the error sum of squares. Computationally, it is convenient to store the quantities \( y'y \) and \( W'y \) and obtain the error sum of squares as

\[ E_{ss} = (y'y) - (y'W)(W'W)^{-1}(W'y) \]

While performing the above matrix multiplication, the resulting product \( \hat{\beta} = (W'W)^{-1} W'y \) should be stored. This estimate of the parametric vector is generally required for output.

Additionally, for the \( i \)-th model term, the so-called sum of squares due to the hypothesis that the means are equal, can be computed conveniently from the expression
\[ \hat{\theta}_i \left[ (W'W)^{-1} \right]^{-1} \hat{\theta}_i \]

where \( \hat{\theta}_i \) is the estimate previously obtained for parameters corresponding to the i-th model term, and \( (W'W)^{-1} \) is the block diagonal submatrix of \( (W'W)^{-1} \) corresponding to the i-th model term. The appropriate positions have been defined in the entries of \( \text{ELW} \) and \( \text{LHW} \), respectively.

Consider now the case in which one or more combinations of levels of the factors is unobserved. The vector \( SX \) was constructed such that the i-th element of \( SX \) contains the sum of the elements of the i-th column of the \( X \) matrix. Therefore, this vector \( SX \) contains zeroes in the positions corresponding to empty cells.

It is necessary to distinguish two situations. If the \( X \) matrix corresponds to a model containing all the admissible interactions, then the information on empty cells contained in \( SX \) is complete. If, however, one or more possible interactions are not contained in the model, then the information on empty cells is not available in \( SX \).

For either case, AARDVARK contains an algorithm which inspects the locations of zeroes in \( SX \) and reparametrizes to full rank using this information. The algorithm is now described and examples are given illustrating the situations for which it is applicable, and situations in which this algorithm leads to erroneous analysis of data.

Consider the case of a two factor classification specified with interaction in which the observations arranged lexicographically
The resulting vector $S_X$ is given by

$$
S_X = \begin{bmatrix}
\mu a_1(1) & a_1(2) & a_1(3) & a_2(1) & a_2(2) & a_2(3) & a_{12}(1,1) & a_{12}(1,2) & a_{12}(1,3) \\
3 & 2 & 4 & 5 & 2 & 2 & 0 & 0 & 0 \\
3 & 0 & 0 & 0 & 2 & 2
\end{bmatrix}
$$

The elements of the vector contain the number of observations which contain, in their expectation, the corresponding parameter. For convenience, the parameters are listed above their incidence. The algorithm proceeds by an inspection of the vector $S_X$ for zeroes and in so doing utilizes the control vectors previously defined.

If a zero is found in a position of $S_X$ corresponding to a main effect, the level of the factor and any interaction parameter containing the factor at this level are deleted. If there remain at least two levels of a factor, the factor listed as "estimable". The control vectors are modified accordingly.
If a zero is found in a position of SX corresponding to an interaction, the corresponding interaction parameter is deleted. Consider the set of indices of the right-most bracket. If there remain at least two levels of the last index for every combination of the remaining indices, the interaction is listed as "estimable".

In the example, the vector modified by the elimination of zeroes would appear as

\[
\begin{bmatrix}
\mu & a_1(1) & a_1(2) & a_1(3) & a_2(1) & a_2(2) & a_2(3) & a_{12}(1,1) & a_{12}(2,1) \\
9 & 2 & 3 & 4 & 5 & 2 & 2 & 3, & 2, & 2 & 2, & 3, & a_{12}(3,2) & a_{12}(3,3) & 2, & 2
\end{bmatrix}
\]

Note that this algorithm has imposed restrictions

\[
a_{12}(1,2) = a_{12}(1,3) = a_{12}(2,2) = a_{12}(2,3) = a_{12}(3,1) = 0.
\]

The above restrictions are on independent functions not estimable with the available data and are therefore valid. Obviously, the data contain no information on these parameters, and therefore any parametric function containing them is non-estimable.

In the example, the interaction model term would be listed as "non-estimable" and therefore deleted. The analysis of variance provided by AARDVARK would result in a decomposition of degrees of freedom given by
This analysis is incorrect for the following reasons. First, there obviously exist five independent comparisons of the observations having zero expectation. Therefore, there should be at least five degrees of freedom for error. In the parameters $a_1(j)$ there exists only one estimable function, and therefore, the data contain only one degree of freedom for $A$. In the parameters $a_2(j)$ there exists only one estimable function, and therefore the data contains only one degree of freedom for $B$.

However, there are also situations in which the algorithm described above leads to a correct analysis. For example if the incidence of observations, in the preceding example, had resulted in an "estimable" interaction, then the resulting analysis would have been correct. Failure of the algorithm to yield the correct analysis in many situations lead this author to recommend its use only in situations characterized by unbalanced complete data.

C. Bradley's Algorithm

Bradley (3) described schemes for finding a set of independent estimable parametric functions for arbitrary arrangements of
observations in factorial designs. Four situations were considered, the scheme being modified in accordance with the situation. These are:

1) No missing cells and a model containing all possible interactions,
2) No missing cells and a model not containing all the possible interactions,
3) Missing cells and a model containing all possible interactions,
4) Missing cells and a model not containing all the possible interactions.

To illustrate the most general scheme, applicable to all the above situations consider the three way linear additive model

\[ y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + e_{ijk} \quad \text{for } i,j,k = 1, 2, 3. \]

Suppose the observed combinations and resulting design matrix, ignoring cell replications, are given by

<table>
<thead>
<tr>
<th>Cell Number</th>
<th>Associated Indices</th>
<th>X</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>1 1 1</td>
<td>1 1 0 0 1 0 0 1 0 0</td>
</tr>
<tr>
<td>(2)</td>
<td>1 2 2</td>
<td>1 1 0 0 0 1 0 0 1 0</td>
</tr>
<tr>
<td>(3)</td>
<td>1 3 3</td>
<td>1 1 0 0 0 0 1 0 0 1</td>
</tr>
<tr>
<td>(4)</td>
<td>2 1 2</td>
<td>1 0 1 0 1 0 0 0 1 0</td>
</tr>
<tr>
<td>(5)</td>
<td>2 2 3</td>
<td>1 0 1 0 0 1 0 0 0 1</td>
</tr>
<tr>
<td>(6)</td>
<td>3 1 3</td>
<td>1 0 0 1 1 0 0 0 0 1</td>
</tr>
</tbody>
</table>

Bradley applies the following procedure.
Rule 1 If the first non-empty cell in the lexicographic ordering of indices is cell designated by \((i_0^*, j_0^*, k_0^*)\), then form a matrix \(W\) by deleting from \(X\) all the columns corresponding to parameters containing \(i = i_0^*, \, j = j_0^*\) or \(k = k_0^*\).

Rule 1 is equivalent to a reparametrization of the linear model, which reduces to full rank the class of models for cases 4.1 and 4.2. Such reparametrization, as defined by the first rule, has an obvious advantage in incomplete models over the conventional procedure of deleting the last level of each factor and associated interaction parameters containing factors at their highest levels. Namely, this rule obviates the possibility of imposing a restriction on the solution that eliminates a combination of the factors that was not observed.

For the example here considered, the resulting matrix \(W\) is given by

\[
\begin{pmatrix}
\mu & \alpha_2 & \alpha_3 & \beta_2 & \beta_3 & \gamma_2 & \gamma_3 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 1 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 \\
\end{pmatrix}
\]

In general, for incomplete factorials, the matrix \(W\) may be deficient in rank and will not provide the required basis for the space of all estimable parametric functions. Bradley's procedure is to apply
the following additional rules.

Rule 2 Form $W^*$ from $W$, rearranging the rows in non-decreasing order of magnitude of the number of 1's in each row.

Rule 3 Mark the uppermost non-zero element in each column of $W^*$, subject to the restriction that no row be marked twice.

Rule 4 Take each unmarked column and eliminate non-zero terms using elementary combinations of marked columns until either the column can be marked by applying Rule 3, or until the column is eliminated. The resulting matrix $Z$ yields a basis for the space of estimable functions.

Concluding the previous example, we note that $W^*$ is given by

$$
\begin{array}{ccccccc}
\mu & \alpha_2 & \alpha_3 & \beta_2 & \beta_3 & \gamma_2 & \gamma_3 \\
1^\# & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1^\# & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1^\# & 0 & 1 \\
1 & 1^\# & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1^\# & 0 & 0 & 0 & 1 \\
1 & 1 & 0 & 1 & 0 & 0 & 1^\# \\
\end{array}
$$

The matrix resulting from the application of Rule 4 with columns rearranged is given by

$$
\begin{array}{ccccccc}
\mu & \beta_2 & \beta_3 & \alpha_2 & \alpha_3 & \gamma_2 & \gamma_3 \\
1^\# & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1^\# & 0 & 0 & 0 & 0 & 0 \\
Z = 1 & 0 & 1^\# & 0 & 0 & 0 & 1
\end{array}
$$
Hence, the Bradley procedure has implicitly imposed $\alpha_1 = \beta_1 = \gamma_1 = \gamma_2 = 0$, and established the rank of the design matrix while providing a basis for the space of estimable functions.

The scheme defined by Bradley has the following important limitations. First, the set of independent estimable functions obtained in general is not the set of functions of general interest to an experimenter. For example, with the cells observed in the preceding experiment, a general computing algorithm should render the following information on estimability:

(i) no simple contrast of the form
\[ \alpha_i - \alpha_j \quad \text{or} \quad \beta_i - \beta_j \quad \text{or} \quad \gamma_i - \gamma_j \]

is estimable,

(ii) \[ E(y_{122} - y_{212} - y_{223} + y_{313}) = \alpha_1 - 2\alpha_2 + \alpha_3, \]
\[ E(y_{122} + y_{133} + y_{212} - y_{223}) = \beta_1 - 2\beta_2 + \beta_3 \quad \text{and} \]
\[ E(y_{111} - y_{122} - y_{212} + y_{223}) = \gamma_1 - 2\gamma_2 - \gamma_3, \]

(iii) no other contrast in the $\alpha$ parameters alone or $\beta$ parameters alone or $\gamma$ parameters alone is estimable.

Secondly, the "marking rule" defined by Bradley requires the computer to "take in turn each column that has no marked element, eliminate terms by adding and subtracting marked columns, until the uppermost
coefficient appears in an unmarked row, or until all coefficients are zero". Such instruction does not describe a programmable algorithm because, clearly, in an experiment of relatively large size, there would exist a very large number of possible operations. There is no attempt made by Bradley to define a column operation sequence guaranteed to terminate successfully. Hence, the method does not provide directly programmable algorithm for an arbitrary number of factors.

D. Methods Using Conditional Inverses

Consider the admissible sets corresponding to any given structure. It is possible to partition any classification model

\[ y = X\beta + e, \]

into

\[ y = \sum_{i=1}^{m} x_i \beta_i + e, \]

(4.18)

where \( x_i \) and \( \beta_i \) are the coefficient matrix and parametric vector corresponding to the \( i \)-th admissible set.

Suppose one is interested in determining the functions of parameters in \( \beta \) alone that are estimable with the observations. Let

\[ z_j = \begin{bmatrix} x_1 & x_2 & \cdots & x_{j-1} & x_{j+1} & \cdots & x_m \end{bmatrix}, \]

\[ \delta_j = (\beta_1', \beta_2', \cdots, \beta_{j-1}', \beta_{j+1}', \cdots, \beta_m'). \]

Then the model in (4.18) may be rewritten as

\[ y = z_j \delta_j + x_j \beta_j + e. \]
The reduced normal equations in $\beta_j$ alone are $X_j'(I-M_j)X_j\beta_j = X_j'(I-M_j)y$, where $M_j$ is the unique orthogonal projection operator on the column space $C(z_j)$ of $z_j$. The matrix $M_j$ is expressible as $z_j(z_j'z_j)^*z_j'$, where $(z_j'z_j)^*$ is any conditional inverse of $(z_j'z_j)$, i.e., $(z_j'z_j)^*$ satisfies $(z_j'z_j)(z_j'z_j)^*(z_j'z_j) = (z_j'z_j)$.

By well known properties of the normal equations

$$(4.19) \quad \mathcal{R}[(I-M_j)X_j] ,$$

where $\mathcal{R}$ denotes the row space, is a basis for the space of all estimable functions in $\beta_j$ alone.

If the elements of $\beta_j$ are parameters corresponding to a main effect and the model contains interactions involving the $\beta_j$, then $C(z_j) \supset C(X_j)$. Hence, in this case, the row space defined in (4.19) is the null space.

However, if one first imposes on the model (4.18) the restrictions given by (4.3), the problem is resolved by obtaining the matrices in (4.19). It should be noted that the known methods of obtaining conditional inverses require a knowledge of the rank of the matrix, or require exact arithmetic operations, or are recursive methods converging on an inverse. Further, if the model contains many admissible sets, the computations required to obtain $(I-M_j)$ for each $j$ are considerable lengthy.

Consider now the problem of determining if the model (4.18) permits the unbiased estimation of a parametric function $\lambda'\beta_i$. A condition for estimability is the consistency of the equations $X'a = \lambda$. A necessary and sufficient condition for consistency is that $r(X') = r(X'_i, \lambda)$. Verification of this condition is tedious because the determination of the rank of an arbitrary matrix is difficult on a digital computer. The round-off
error inherent in most procedures at times precludes exact determination of rank.

Rao (22) states that a conditional inverse \( X'^* \) of maximum rank can be used to test the consistency of the given equations. The theorem stated is as follows. A necessary and sufficient condition for the consistency of the equations \( X'a = \lambda \) is that the \( i \)-th component of \( X'^*\lambda \) be zero when the \( i \)-th row of \( X'^*X' \) is zero, where \( X'^* \) has maximum rank.

The condition given by Rao is necessary because if \( X'a = \lambda \) are consistent, then there exists \( a_0 \) such that \( X'a_0 = \lambda \), and thus \( X'^*X'a_0 = X'^*\lambda \). It follows that the \( i \)-th component of \( X'^*\lambda \) is zero whenever the \( i \)-th row of \( X'^*X' \) is zero.

That Rao's condition is not sufficient is shown in the following counterexample. The matrix \( J \) of order \( n \times m \), with \( n < m \), each of whose elements is 1, has a conditional inverse

\[
J^* = \frac{1}{n} \begin{pmatrix} 1 \\ 0^T \end{pmatrix}
\]

The rank of \( J^* \) is \( n \) and therefore \( J^* \) is of maximum rank. The \( i \)-th row of \( J^*J \) equals zero only if \( i > n \), and for \( i > n \) the \( i \)-th element of \( J^*\lambda \) is zero for any \( \lambda \). However, \( Jx = \lambda \) are not consistent for arbitrary \( \lambda \). Hence, the condition is not sufficient for consistency.

It seems interesting to note that this author has not been able to locate any computer program for determining the consistency of an arbitrary set of equations. If one assumes exact arithmetic, the following known facts lead to a programmable consistency test for any arbitrary coefficient matrix \( A \).
Every rectangular matrix $A$ of rank $r$ is row equivalent to a matrix $BA$ of echelon form, defined as follows:

(a) the first $r$ rows are non-zero and all remaining rows, if any, are zero,

(b) in the $i$th row ($i = 1, 2, \ldots, r$), the first non-zero element is unity, the column in which it occurs being labelled $C_i$,

(c) $C_1 < C_2 < C_3 < \ldots < C_r$,

(d) in the columns $C_i$, the only non-zero element is the 1 in row $i$.

**Lemma 4.1**

A necessary and sufficient condition for consistency of the equations $Ax = \lambda$ is that the $i$th element of $B\lambda$ be zero whenever the $i$th row of $BA$ is zero, where $B$ is any non-singular matrix such that $BA$ is in echelon form.

The equations $Ax = \lambda$ are equivalent to

$$BAX = B\lambda .$$

There exist a $C$ that is non-singular and permutes the columns of $BA$ such that

$$BACy = B\lambda ,$$

where

$$BAC = \begin{bmatrix}
I_r & D \\
0 & 0
\end{bmatrix}, \quad \text{and} \quad Cy = x .$$

It is clear that the equations in (4.20) are consistent if and only if
for every $i > r$, the $i$-th element of $B\lambda$ is zero.

A difficulty in applying the test for consistency specified Lemma 4.1 is the requirement of exact arithmetic operations in order to identify zeroes.

The simplex algorithm for solution of linear programming problems can be used effectively to determine consistency of any set of equations. The classical linear programming problem maximizes a linear function $c'x$ subject to a set of linear constraints $Ax \leq b$, where all components of the vector $x$ are nonnegative.

Any computer installation having capabilities for solving a linear programming problem can use this same program to determine the consistency of a set of equations $Ax = b$, where $A$ is any rectangular matrix. At the same time, the procedure described provides a solution to the set of equations in the event that they are consistent.

Under all circumstances $x_i$, the $i$-th component of $x$, can be expressed as $x_i = z_i - z_{n+i}$, $i = 1, 2, \ldots, n$, where all components of the $2n \times 1$ vector $z$ of $z_i$'s are nonnegative. Thus, $x = (1_n', -1_n)z$ and $Ax = (A, -A)z = Bz$. Hence, whenever the equations $Ax = b$ are consistent then there exists a vector $z$ of nonnegative elements such that $(A, -A)z = b$. This fact can be put in a linear programming context as follows.

**Lemma 4.2**

A set of equations $Ax = b$ are consistent if and only if the linear programming problem maximizing $c'z$ subject to $(A, -A)z = b$, with $z \geq 0$, has a feasible solution.
A convenient choice of $c$ in the objective function $c'z$ is any one row of the matrix $B = (A, -A)$.

If, in particular, $b'\beta$ is a parametric function of interest and $A$ is the matrix $X'X$, coefficient matrix of the conjugate normal equations $X'Xx = b$, then a solution vector $z$ of the corresponding linear programming problem transformed to the original variables in $x$ yields the best linear unbiased estimate of $b'\beta$ given by $x'X'y$.

Alternatively, one may impose the usual restrictions (4.3) on the model in (4.18). The resulting normal equations

$$W'W\theta = W'y,$$

may or may not have a unique solution. If $(W'W)^*$ is any conditional inverse of $W'W$, then $W(W'W)^*W'$ is the orthogonal projection operator on the column space of $W$. If $\lambda'\theta$ is estimable then the equations $W'a = \lambda$ are consistent in the vector $a$. Hence,

$$\lambda'(W'W)^*W'W = a'(W'W)^*W'W = a'W = \lambda',$$

and thus estimability of $\lambda'\theta$ implies $\lambda'H = \lambda'$, where $H = (W'W)^*W'W$.

Conversely, if $\lambda'H = \lambda'$, then the equations $\lambda' = a'W$ are satisfied with $a' = \lambda'(W'W)^*W'$. Thus, $\lambda'\theta$ is estimable if and only if $\lambda'H = \lambda'$. This result is equivalent to the one stated by Rao (22) that $\lambda'\theta$ is estimable if and only if $\lambda'(I-H) = 0$.

If the experimental arrangement is of maximal rank, the conditions imposed in (4.3) are necessary and sufficient for $H = I$. If $H \neq I$, it follows that $\theta_i$ is estimable if and only if $h_i = e_i$, where $h_i$ is the $i$-th row of $H$ and $e_i$ is the $i$-th row of the identity matrix $I$. 
Hence, by inspection of the H matrix one can immediately determine which \( \theta_i \) are estimable. By using the imposed restrictions, any estimable \( \theta_i \) can be expressed as a function of the original parameters.

It should be clear that while the condition \( h_i = e_i \) is necessary and sufficient for \( \theta_i \) to be estimable, other functions of \( \theta \) may be estimable and these require verification of the invariance \( \lambda' H = \lambda' \).

E. Wilkinson's Recursive Algorithm for A.O. V.

Wilkinson (32) outlined a computational recursive algorithm for the analysis of experimental design data. This algorithm is based on one simple operation termed a sweep. A sweep is defined as the computation of a mean or means and the subtraction of these means from a prescribed data set. The claim made by Wilkinson is that "any experimental design however complex can be analyzed with a finite sequence of sweep operations". Hence, the analysis of variance can be performed by defining a control vector to direct the order of the sweeps and the elements from which each mean will be subtracted.

This section discusses the method and comments on some aspects of its applicability for the analysis of data on a digital computer.

The method proposed by Wilkinson is compared to that described in Chapter II for balanced data and that described in Section B for non-orthogonal analysis.

The theoretical basis for the Wilkinson algorithm consists of the
step-wise model fitting of each factor of classification accompanied by a readjustment of the fit after each step. However, the recursive procedure is not a direct application of this theory. The examples presented by Wilkinson certainly indicate that the recursive algorithm leads to correct analysis for certain incomplete block designs. Balanced complete block designs can also be analysed by a sequence of sweeps. An example is presented in this section. However, it is also shown that this method does not lead to the correct sums of squares for arbitrarily incomplete factorial arrangements.

Described briefly, the method of Wilkinson first requires an ordering of the classification factors in an experiment. When this method is applied to arbitrarily incomplete experimental data, the analysis it produces is dependent on the order in which the factors are arranged.

The symbol $S_i$ will denote the residual vector after the $i$-th sweep. $S_1$ will always be the residual vector after fitting for the mean. $S_i$ is computed from the current residual vector by obtaining the admissible partial means corresponding to the $i$-th factor of classification and then subtracting, from each element in the current residual vector, the mean to which that element contributed. Because a factor of classification defines a partitioning of the current residual vector into disjoint exhaustive sets, it follows that each element enters into the computation of one and only partial mean.

For example, consider the balanced complete two factor additive model

$$y_{j_1j_2} = \mu + a_1(j_1) + a_2(j_2) + e_{j_1j_2},$$

where
\[ j_1, j_2 = 1, 2. \]

By arranging the observations lexicographically, \( S_1 \) is obtained as the fourth column of the following display.

<table>
<thead>
<tr>
<th>Observation</th>
<th>( j_1 )</th>
<th>( j_2 )</th>
<th>( y_{j_1 j_2} )</th>
<th>( S_1 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td></td>
<td>-1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>2</td>
<td></td>
<td>-2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>5</td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>5</td>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

\[ S_{1j_{1j_2}} = y_{j_{1j_2}} - \bar{y} \cdot, \]

is the residual after fitting the model

\[ y_{j_{1j_2}} = \mu + e_{j_{1j_2}}. \]

Then, \( \mu \) is estimated as 4.

Next, \( S_2 \) is obtained as follows:

<table>
<thead>
<tr>
<th>( j_1 )</th>
<th>( j_2 )</th>
<th>( S_1 )</th>
<th>( S_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>( \frac{1}{3} )</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>-2</td>
<td>-( \frac{1}{3} )</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>( \frac{2}{3} )</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>1</td>
<td>( \frac{1}{3} )</td>
</tr>
</tbody>
</table>

\[ S_2 = (y_{j_{1j_2}} - \bar{y} \cdot) - \sum_{j_2} (y_{j_{1j_2}} - \bar{y} \cdot)/2, \]

is the residual vector after fitting the model.
The unadjusted effects are estimated by
\[ S_{j_1 j_2} = a_1(j_1) + e_{j_1 j_2}. \]

The unadjusted \( a_1(j_1) \) effects are estimated by
\[ \sum_{j_2} (y_{j_1 j_2} - y_{..})/2. \]

Then, the estimate of \( a_1(1) \) is \(-3/2\), while \( a_1(2) \) is estimated as \( 3/2 \).

Next, \( S_2 \) is obtained as follows:

\[
\begin{array}{cccc}
  j_1 & j_2 & S_2 & S_3 \\
  1 & 1 & \frac{1}{2} & 0 \\
  1 & 2 & -\frac{1}{2} & 0 \\
  2 & 1 & \frac{1}{2} & \rightarrow 0 \\
  2 & 2 & -\frac{1}{2} & 0 \\
\end{array}
\]

\[ S_{j_1 j_2} = (y_{j_1 j_2} - y_{..}) - \sum_{j_2} (y_{j_1 j_2} - y_{..})/2 - \left( \sum_{j_1} \left[ (y_{j_1 j_2} - y_{..}) - \sum_{j_2} (y_{j_1 j_2} - y_{..})/2 \right] \right)/2, \]

is the residual vector after fitting the model

\[ S_{j_1 j_2} = a_2(j_2) + e_{j_1 j_2}. \]

The unadjusted \( a_2(j_2) \) effects are estimated by
\[ \left( \sum_{j_1} \left[ (y_{j_1 j_2} - y_{..}) - \sum_{j_2} (y_{j_1 j_2} - y_{..})/2 \right] \right)/2. \]

Then, the estimate of \( a_2(1) \) is \( \frac{1}{2} \), while \( a_2(2) \) is estimated as \( -\frac{1}{2} \).
Because the residual vector after fitting the $a_2(j)$ parameters is null, no further sweeps or readjustments are required. This, Wilkinson termed first order balance.

The observation vector has been decomposed orthogonally as

$$y_{1j_2} = \mu + a_1(j_1) + a_2(j_2) + e_{1j_2}$$

$$\begin{bmatrix} 3 \\ 2 \\ 6 \\ 5 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \\ 4 \\ 4 \end{bmatrix} + \begin{bmatrix} -3/2 \\ -3/2 \\ +3/2 \\ +3/2 \end{bmatrix} + \begin{bmatrix} 1/2 \\ -1/2 \\ 1/2 \\ -1/2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

The corresponding analysis of variance table would have entries for sums of squares and degrees of freedom, given by

Total = $\mu + a_1 + a_2 + e$

S.S. 74 = 64 + 9 + 1 + 0

D.F. 4 = 1 + 1 + 1 + 1.

This essentially, is the Wilkinson recursive algorithm for analysis of variance. The modifications required for incomplete data are given by the following example. At the same time, this next simple example points out (we hope dramatically) that the method cannot be applied to arbitrarily incomplete factorials.

Consider the incomplete but balanced (see Definition 2.10) situation depicted by
where $\#$ indicates unobserved cells, and the numbers in the table correspond to the observations.

A sweep for the mean yields residuals

$$
\begin{array}{|c|c|c|c|}
\hline
\text{Residuals} & 1 & 0 & \#
\hline
\# & 1 & \# & -1
\hline
-1 & \# & 0 & \#
\hline
\# & -1 & \# & 1
\hline
\end{array}
$$

with the estimate of the mean being $\hat{\mu} = 2$.

A sweep for factor $a_1$ is obtained by computing the $a_1$ means and subtracting from the first residual vector. This yields

<table>
<thead>
<tr>
<th>$a_1$</th>
<th>Means</th>
<th>Residuals after fitting $a_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1(1)$</td>
<td>$1/2$</td>
<td>$1/2$ $#$ $-1/2$ $#$</td>
</tr>
<tr>
<td>$a_1(2)$</td>
<td>$0$</td>
<td>$#$ $1$ $#$ $-1$</td>
</tr>
<tr>
<td>$a_1(3)$</td>
<td>$-1/2$</td>
<td>$-1/2$ $#$ $1/2$ $#$</td>
</tr>
<tr>
<td>$a_1(4)$</td>
<td>$0$</td>
<td>$#$ $-1$ $#$ $1$</td>
</tr>
</tbody>
</table>
A sweep for factor $a_2$ is next obtained following the same procedure as above. This yields

<table>
<thead>
<tr>
<th>Means</th>
<th>Residuals after fitting $a_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_2(1)$ 0</td>
<td>$1/2$ $\neq$ $-1/2$ $\neq$</td>
</tr>
<tr>
<td>$a_2(2)$ 0</td>
<td>$\neq$ 1 $\neq$ $-1$</td>
</tr>
<tr>
<td>$a_2(3)$ 0</td>
<td>$-1/2$ $\neq$ 1/2 $\neq$</td>
</tr>
<tr>
<td>$a_2(4)$ 0</td>
<td>$\neq$ $-1$ $\neq$ 1</td>
</tr>
</tbody>
</table>

In obtaining the sweep for the factor $a_2$, Wilkinson divides the $a_2(j_2)$ total by a factor $r$, where $r$ is, as termed by Wilkinson, an effective replication factor. This author admits he was not able to understand exactly how $r$ was obtained, although the author conjectures that $r$ would be $1/2$ in the preceding example. For this reason, the example was designed so that the $r$ value would not affect the results. Clearly, it can not, since the $a_2$ totals are each equal to zero.

In continuing with Wilkinson's analysis, the next step is an adjustment for the $a_1$ factor. This adjustment is made by a re-sweep for factor $a_1$ and yields residuals given by

<table>
<thead>
<tr>
<th>$a_1$ adjustment</th>
<th>Residuals after adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$a_1(1)$ 0</td>
<td>$1/2$ $\neq$ $-1/2$ $\neq$</td>
</tr>
<tr>
<td>$a_1(2)$ 0</td>
<td>$\neq$ 1 $\neq$ $-1$</td>
</tr>
<tr>
<td>$a_1(3)$ 0</td>
<td>$-1/2$ $\neq$ 1/2 $\neq$</td>
</tr>
<tr>
<td>$a_1(4)$ 0</td>
<td>$\neq$ $-1$ $\neq$ 1</td>
</tr>
</tbody>
</table>

Subtracted
The orthogonal decomposition of the observation vector is given by

\[ y_{j_1j_2} = \mu + a_1(j_1) + a_2(j_2) + e_{j_1j_2} \]

The corresponding analysis of variance table would have entries for sums of squares and degrees of freedom, given by

- **Total** = \( \mu + a_1 + a_2 + e \)
- **S.S. 38** = 32 + 1 + 0 + 5
- **D.F. 8** = 1 + 3 + 3 + 1

We now comment critically on the value and validity of the preceding analysis, as well as on the algorithm in general. On the positive side, the algorithm does accomplish an orthogonal additive decomposition of the observational vector. However, it is clear from an inspection of the arrangement, that the design does not have three degrees of freedom for each of the main effects. There are, in fact, only two linearly independent estimable functions within each factor of classification, and therefore, only 2 degrees of freedom for each factor. Further, without special assumptions, the situation represented by the arrangement has no
degrees of freedom for error. That is, if the additive two factor model is assumed, then the above arrangement yields no linear independent functions of the observation which estimate zero unbiasedly, and these generally correspond to individual degrees of freedom for error.

Although no formal attempt has been made to identify the classes of situations to which the Wilkinson procedure applies, it is clear from Wilkinson's examples that the algorithm is useful for certain incomplete block designs. These classes of data situations are characterized as having maximal rank in each of the factors of classification, thus avoiding the difficulties elucidated in the example.

The algorithm, however, has additional disadvantages, from this author's point of view. These are:

1) The results of the analysis are dependent on the order of presentation of the factors.

2) The algorithm does not provide information on degrees of freedom for situations of less than maximal rank.

3) From the examples given by Wilkinson, and the description of the algorithm, it would seem to be applicable to two factor arrangements only. It is not clear, at least to this author, how the algorithm would proceed to analyse incomplete models containing three or more factors of classification. It is possible, however, that Wilkinson can describe the necessary modification required for an arbitrary number of factors, and this would certainly make the algorithm more valuable and generally
more useful.

F. Contribution of Elston and Bush

Elston and Bush (8) defined a set of hypotheses testable when there are interactions in an analysis of variance model. Their intent was to identify hypotheses about the main effects and to provide a method of obtaining the sum of squares appropriate for testing any testable hypothesis. In particular, the authors considered the problem of determining what is testable when one or more subclasses are empty. They explicitly dealt only with two-way arrangements and concluded that "the principles for a higher-way classification are exactly the same". Because the work by Elston and Bush motivated the study contained in Chapter V, their results will be briefly presented here along with a few pertinent comments. It is desirable to point out that the following is concerned only with the case of empty subclasses, and that the reference (8) contains additional results not covered in this section.

The authors state that it is possible to "develop testable hypotheses that test for the main effects and interactions to the extent that the data allow...", even when one or more subclasses are empty. In the case of just one empty subclass with mean $\mu_{pq}$ (i.e., the $p$-th and $q$-th levels of the two factors correspond to the empty class), after imposing suitable restrictions, they suggest testing hypotheses of the type

$$H_0 : \sum_j w_j \mu_{i,j} = \sum_j w_j \mu_{i',j},$$

$$i \neq i'; i, i' = 1, 2, \ldots, p - 1, p + 1, \ldots, a,$$
where \( w_j \) are arbitrary weights, \( \mu_{ij} \) is the mean of the \((i,j)\) subclass, and \( a \) is the number of levels of the first factor. This corresponds to a test for the first factor, leaving out the \( p \)-th level of \( A \). The hypothesis obtained has \( a - 2 \) degrees of freedom. Additionally, they suggest testing

\[
H_0 : \sum_{i \neq p} v_i \sum_{j \neq q} w_j \mu_{ij} = \left( \sum_{i \neq p} v_i \right) \sum_{j \neq q} w_j \mu_{pj},
\]

a hypothesis with one degree of freedom which "can be interpreted as testing whether, when \( w_q = 0 \), \( \sum_j w_j \mu_{pj} \) is equal to the weighted average value of \( \sum_j w_j \mu_{ij} \) for all \( i \neq p \). The preceding hypotheses, jointly, yield a hypothesis with a \( -1 \) degrees of freedom.

The authors extend the preceding to the case of arbitrary number of missing cells and conclude that "it is possible to derive a reasonable hypothesis ... whatever the pattern of empty subclasses, provided only that the filled subclasses form a connected design and at least one level .. (of the factor) .. has no empty subclasses in it". Further, ".. if every level .. (of the factor) .. has at least one empty subclass, then it is still possible to develop a test for the factor with (a-1) degrees of freedom provided the design is connected".

The concept of connectedness is defined and explored in the next chapter. That the connectedness criterion does not usefully generalize to multi-way classifications, as implied by the authors, is demonstrated in Section B of Chapter V.
V. ESTIMATION AND ITS RELATIONSHIP TO CONNECTEDNESS, REDUCIBILITY, AND GRAPHS

A. Introduction

Chapter IV exhibited the problems that arise in the analysis of classification data containing missing cells. Specifically, when the pattern of observed cells forms an arbitrary arrangement, it is desirable to identify the functions within factors of classification that can be estimated.

Bose (2) gave a necessary and sufficient condition for the estimability of every treatment contrast in a block-by-treatment additive model. This condition on the pattern of observed cells he called "connectedness". The literature (8, 23) contains references to this concept as a criterion for estimability, testability, and maximality of rank. As discussed in Section F of Chapter IV, some authors (8') seem to feel that the concept of connectedness could be generalized to multi-way classification.

In this chapter, the major result is a method for identification of the functions that can be estimated within each factor of classification from an arbitrary set of data. The presentation of this algorithm is preceded, in Section B and C, by a series of arguments which represent, essentially, the logical process through which the algorithm was derived. This form of presentation seems justified because of the uncommon nature of the relationships and results employed. In particular, the chapter contains a counterexample that denies the possibility of generalizing Bose's theorem (Section B). Further, the equivalence
between three seemingly unrelated concepts is demonstrated (Section C),
and the relevance of these equivalences to attain computational simplicity
is made manifest in the algorithm described (Section D).

B. Bose's Theorem and its Extension

A statement of Bose's theorem is as follows. A treatment $t_j$ is said
to be associated with a block $b_i$ if there is at least one observation in
the $(i,j)$ subclass. Two treatments, two blocks, or a treatment and a block
are said to be connected if it is possible to pass from one to another by
a chain, such chain consisting alternately of levels of blocks and treat­
ments such that any two adjacent members of the chain are associated. The
block-by-treatment arrangement is said to be connected if every block and
treatment is connected to every other block and treatment. Then Bose (2)
proved the following.

Theorem 5.1

With the model

$$ y_{ijk} = \mu + b_i + t_j + e_{ijk} $$

every treatment difference is estimable if the arrangement is connected.

The connectedness criterion is computationally much easier to verify
than the maximality of the rank of the coefficient matrix. Of course,
Bose's theorem establishes the equivalence of the two. At this point, we
wish to consider the possibility of extending Bose's theorem to three-way
additive models. Such extension would require modification of the pre­
ceding definitions. The following definitions are therefore presented.
Definition 5.1

The $i$-th level of a factor $a_k$, and the $j$-th level of a factor $a_\ell$, are said to be associated if the combination $(a_k(i), a_\ell(j))$ occurs in the data.

Definition 5.2

A chain is a sequence of levels of factors in which any pair of adjacent elements of the chain is associated.

Definition 5.2a

Two levels of two distinct factors are said to be pairwise connected if it is possible to construct a chain containing the two levels and consisting alternately of levels of the two factors.

Definition 5.2b

Two levels of the same factor are said to be pairwise connected with respect to a second factor, if it is possible to construct a chain consisting exclusively of levels of the two factors, containing the two levels of the first factor as members.

Definition 5.3

A factorial arrangement will be said to be pairwise connected if for every two factors, every two levels of any of the factors are pairwise connected.

To clarify these somewhat modified definitions, consider the three factor arrangement whose incidence is depicted by
where 1's correspond to observed subclasses, and 0's to those not observed.

A set of chains which depict all possible associations are:

- $C_1: a_1(1), a_2(1), a_1(2), a_2(2), a_1(1), a_2(3)$
- $C_2: a_1(1), a_3(1), a_1(2), a_3(2), a_3(1), a_3(3)$
- $C_3: a_2(1), a_3(1), a_2(2), a_3(2)$
- $C_4: a_2(3), a_3(3)$

Note that in the arrangement $a_3(3)$ and $a_2(1)$ are not pairwise connected since no chain consisting alternately of associated levels of the two factors can connect these two levels. The arrangement is not pairwise connected.

It is possible to implement on a digital computer the verification of pairwise connectedness of an arrangement. Suppose that for any n-way classification, the program has available the observed combinations for any two of the factors. For example,

<table>
<thead>
<tr>
<th>$a_k(i)$</th>
<th>$a_k(j)$</th>
<th>$a_k(i)$</th>
<th>$a_k(j)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>5</td>
<td>4</td>
</tr>
</tbody>
</table>

Beginning with any level of any of the two factors, say $a_k(1)$, apply the
following rules.

1) Form all sequences of pairs consisting of the starting levels and all of its associates. We have, a) \( a_k(1), a_{\ell}(1) \), b) \( a_k(1), a_{\ell}(2) \).

2) To the last entry of each sequence, reapply rule 1 subject to the restriction that a level of the same factor not appear twice in the same sequence. We now have, a) \( a_k(1), a_{\ell}(1) \), b) \( a_k(1), a_{\ell}(2), a_k(3) \), c) \( a_k(1), a_{\ell}(2), a_k(5) \).

3) Apply rule 2 until every sequence terminates. This gives a) \( a_k(1), a_{\ell}(1) \), b) \( a_k(1), a_{\ell}(2), a_k(3), a_{\ell}(3), a_k(4) \), c) \( a_k(1), a_{\ell}(2), a_k(5), a_{\ell}(4), a_k(2) \).

4) If the resulting sequences contain all the levels of each of the factors, as in the above example, then the factors are pairwise connected. Otherwise, they are not.

We can immediately verify that these modified definitions do not affect the validity of Bose's theorem. The following is a modified statement and proof of Bose's theorem within the context of definitions 5.1 through 5.3.

**Theorem 5.2**

If every level of each factor is observed in a model

\[ y_{ijk} = \mu + a_{1}(i) + a_{2}(j) + e_{ijk}, \quad E(e_{ijk}) = 0 \quad \text{for all } i, j, k, \]

where \( i = 1, 2, \ldots, r_{1}, \quad j = 1, 2, \ldots, r_{2} \quad \text{and } k = 0, 1, 2, \ldots, n_{ij} \), the pattern of observed cells forms a pairwise connected design if and only if every contrast,
\[ \sum_{i=1}^{r_1} \lambda_i a_1(i), \text{ where } \sum_{i=1}^{r_1} \lambda_i = 0 \]

is estimable.

**Proof:** If the arrangement is connected, all \( r \) levels of \( a_1 \) are connected, and therefore there exists a chain \( C_1 \) such that

\[ C_1: a_1(i_1), a_2(j_1), a_1(i_2), a_2(j_2), \ldots, a_2(j_p-1), a_1(i_p) \]

contains all the levels of \( a_1 \), and every pair of adjacent members is associated. But,

\[ \sum_{i=1}^{r_1} \lambda_i a_1(i) = \sum_{i=1}^{r_1} \lambda_i (a_1(i) - a_1(1)) \]

\[ = \sum_{i=1}^{r_1-1} \lambda_i (a_1(i) - a_1(1)), \]

because

\[ \sum_{i=1}^{r_1} \lambda_i a_1(1) = 0. \]

For each \( i = 1, 2, \ldots, r_1-1 \), \( a_1(i) - a_1(1) \) can be estimated unbiasedly as follows. Let \( C_i \) be that segment of the chain \( C_1 \) that connects \( a_1(i) \) to \( a_1(1) \). Then for each \( i \)

\[ C_i: a_1(i) = a_1(i_1), a_2(j_1), a_1(i_2), a_2(j_2), \ldots, a_2(j_{m+1}) = a_1(1) \]
If \( y(i,j) \) denotes an observation in the \((i,j)\) cell, \( C_i \) defines an unbiased estimator of \( a(i) - a_(r_1) \) given by

\[
(5.3) \quad y(i_m^, j_m^) - y(i_{m+1}^, j_m^) + y(i_{m+1}^, j_{m+1}^) - y(i_{m+2}^, j_{m+1}^) + \ldots
- y(i_{m+p}^, j_{m+p-1}^).
\]

Since \( C_i \) exists for all \( i \), it follows that all \( a(i) - a_(r_1) \) are estimable, and therefore every contrast is estimable.

Conversely, if \( \sum_{i=1}^{r_1} l_i a(i) \) is estimable for all \( \ell_i \) satisfying \( \sum_{i=1}^{r_1} l_i = 0 \), it follows that \( a(i') - a_(r_1) \) is estimable for each \( i' \).

If \( E[\sum_{i,j,k} \lambda_{ijk} y_{ijk}] = a(i') - a_(r_1) \), it follows that \( \sum_{i,j,k} \lambda_{ijk} = 0 \) for each fixed value of \( j \); \( \sum_{j,k} \lambda_{i',jk} = 1 \); \( \sum_{j,k} \lambda_{i,j,k} = -1 \); and \( \sum_{j,k} \lambda_{ijk} = 0 \) for \( i \neq i' \neq r_1 \). The observations in the estimator \( \sum_{i,j,k} \lambda_{ijk} y_{ijk} \) having coefficients distinct from zero can be arranged in the form of expression 5.3, thereby defining a chain, as in 5.2. The existence of partial chains for \( i' = 1, 2, \ldots, r_1 \) implies the existence of the chain 5.1 and therefore the connectedness of the design.

We now turn our attention to the possibility of extending Theorem 5.2 to a three-way additive classificatory arrangement. That pairwise connectedness does not imply estimability of all contrasts within a factor of classification is immediately obvious from the following counterexample.
Consider the three-way classification model

\[ y_{ijkl} = \mu + a_1(i) + a_2(j) + a_3(k) + e_{ijkl}, \quad \mathbb{E}(e_{ijkl}) = 0 \]

with \( i, j, k = 1, 2, 3 \) and \( \lambda = 0, 1, 2, \ldots, m_{ijk} \).

Suppose six of the 27 cells were observed, and these are given by indices

<table>
<thead>
<tr>
<th>Cell No</th>
<th>i</th>
<th>j</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(2)</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(3)</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>(4)</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(5)</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>(6)</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

The following chains show respectively that \((a_1(i), a_2(j), a_3(k))\), and \((a_1(i), a_3(k))\) are pairwise connected.

\[ C_1 : a_1(1), a_2(2), a_2(2), a_1(1), a_3(3), a_1(1), a_2(1), a_1(3) \]
\[ C_2 : a_2(1), a_3(1), a_2(1), a_3(2), a_2(2), a_3(3), a_2(3), \]
\[ C_3 : a_1(1), a_3(1), a_1(1), a_3(2), a_1(2), a_3(3), a_1(3). \]

Therefore, the existence of the above chains shows that the arrangement is pairwise connected. However, no simple contrasts \([a_m(i) - a_m(i')]\) are estimable for any value of \( m, i \) and \( i', i \neq i' \). Hence, for the given model pairwise connectedness is not a sufficient condition for estimability of all linear main effect contrasts.
While the counterexample shows that Bose's theorem does not generalize, the following theorem states that pairwise connectedness is a necessary condition for estimability of all main effect contrasts in any additive model.

Theorem 5.3

With the model

\[ y_i = \mu + \sum_{s=1}^{n} a_s(i_s) + e_i; \quad E(e_i) = 0; \quad i_s = 1, 2, \ldots, r_s \]

estimability of all contrasts of the form,

\[ \sum_{i=1}^{r_s} f_i a_s(i) \quad \text{with} \quad \sum_{i=1}^{r_s} f_i = 0 \quad \text{for every} \quad s \]

implies the arrangement forms a pairwise connected design.

Proof If the arrangement is not pairwise connected, by Definition 5.3, there exists a pair of factors \( a_p, a_q \), \( p \) not necessarily different from \( q \), and two levels of these factors, \( a_p(i), a_q(j) \), such that the chain of Definition 5.2 cannot be constructed. Without loss of generality, assume there exists a chain connecting every level of \( a_p(i') \), \( i \neq i' \), to every level \( a_q(j') \), \( j \neq j' \). Then \( a_p(i) - a_p(i') \) is non-estimable since the set of levels of \( a_q(i) \) associated with \( a_p(i) \) are disjoint from that set of levels of \( a_q(i) \) associated with \( a_p(i') \).

Having observed the impossibility of generalizing Bose's theorem, the author's attention was drawn to the relationship between the concept of connectedness and the mathematical concepts of irreducibility.
of matrices and graphs associated with matrices.

C. The Equivalence Between Connectedness, Irreducibility of the Incidence Matrix, and a Property of its Graph

We first demonstrate the equivalence between connectedness and irreducibility. The following definitions are useful.

**Definition 5.4**

For any given arrangement of observed cells on two factors of classification $a_k$, $a_\ell$, with $r_k$ and $r_\ell$ levels respectively, the $r_k \times r_\ell$ matrix $N_{k,\ell}$ defined by

$$N_{k,\ell}(i,j) = \begin{cases} 1 & \text{if the } i\text{-th level of } a_k \text{ is associated with the } j\text{-th level of } a_\ell, \\ 0 & \text{otherwise,} \end{cases}$$

will be called the association matrix of the two factors.

**Definition 5.5**

A square matrix $A$ of order $n$ is called reducible if the index set $I_n = \{1, 2, \ldots, n\}$ can be split into two complementary disjoint sets $I_i = \{i_1, i_2, \ldots, i_d\}$ and $I_k = \{k_1, k_2, \ldots, k_d\}$, such that

$$A(i_\alpha, k_\beta) = 0 \quad \text{for all pairs } (i_\alpha, k_\beta) \text{ such that } i_\alpha \in I_i \text{ and } k_\beta \in I_k.$$
Lemma 5.1

If $N_{k,k}$ is the association matrix of factors $a_k$ and $a_k$ respectively, then $N_{k,k} = N_{k,k}^*$

Lemma 5.2

The association matrix of a factor with itself $N_{k,k}$ is diagonal and the $i$-th diagonal element is equal to one if the $i$-th level of the factor was observed, equal to zero otherwise.

Lemma 5.3

If $N_k = N_{k,k} N_{k,k}^*$, then

$N_k^k(i,i) = \text{the number of levels of factor } a_k \text{ associated with the } i\text{-th level of } a_k.$

$N_k^k(i,j) = \text{the number of levels of factor } a_k \text{ associated with both the } i\text{-th level and } j\text{-th levels of } a_k.$

Lemma 5.4

A square matrix $A$ is reducible if there exists a permutation matrix $P$ such that

$P A P^T = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix}$

where $B$ and $D$ are square matrices and $0$ is a zero matrix.

Theorem 5.4

For the model
\[ y_{ijk} = \mu + a_1(i) + a_2(j) + \epsilon_{ijk}, \quad \mathbb{E}(\epsilon_{ijk}) = 0 \]

with \( i = 1, 2, \ldots, r_1, \quad j = 1, 2, \ldots, r_2, \quad k = 0, 1, \ldots, n_{ij}, \)

irreducibility of \( N_{ij}^1 = N_{12}^1 N_{21}^1 \) implies and is implied by estimability of all contrasts \( a_1(i) - a_1(i') \).

**Proof** Any off-diagonal of \( N_{ij}^1 \), such as \( N_{ij}^1(i, i') \), is greater than zero if and only if \( a_1(i) - a_1(i') \) is estimable. This follows directly from Lemma 5.3, because \( N_{ij}^1(i, i') > 0 \) implies that there are one or more levels of \( a_2 \) that are associated with both \( a_1(i) \) and \( a_1(i') \). If \( N_{ij}^1 \) is reducible, then by Lemma 5.4, there exists a permutation of rows and columns of \( N_{ij}^1 \) such that

\[
P_{i'j'} = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}
\]

with \( A \) and \( B \) square matrices of order \( r_1 - p \) and \( p \), respectively. The dimensions of \( A \) and \( B \) partition the index set \( I_{r_1} \) in accordance with Definition 5.5. Then the functions \( [a_1(i) - a_1(i')] \) are estimable if and only if \( i \) and \( i' \) belong to the same index set. Therefore, estimability of all contrasts \( [a_1(i) - a_1(i')] \) implies the irreducibility of \( N_{ij}^1 \).

Conversely, if \( N_{ij}^1 \) is irreducible, then any contrast \( [a_1(i) - a_1(i')] \) can be estimated as follows.

If \( N_{ij}^1(i, i') > 0 \), the contrast is estimable. If \( N_{ij}^1(i, i') = 0 \), then the irreducibility of \( N_{ij}^1 \) guarantees that for some \( i^\Xi \) such that \( i^\Xi \neq i' \), \( N_{ij}^1(i, i^\Xi) > 0 \). Hence, \( [a_1(i) - a_1(i^\Xi)] \) is estimable. If \( N_{ij}^1(i', i^\Xi) > 0 \), then \( [a_1(i') - a_1(i^\Xi)] \) is estimable and, therefore, so is \( [a_1(i) - a_1(i')] \). If \( N_{ij}^1(i', i^\Xi) = 0 \), the irreducibility of \( N_{ij}^1 \) guarantees
that for some $i, i' \in \mathcal{I}$ such that $i, i' \neq i', N_2(i, i') > 0$, and so on. Hence, irreducibility of $N_2$ implies estimability of all treatment contrasts.

The equivalence between connectedness and irreducibility of the incidence matrix gives us alternative computational methods of verifying the maximality of the rank of the coefficient matrix without actually finding the rank. The desire to verify irreducibility in the simplest possible way led the author to a further equivalence that has interesting ramifications for computational simplicity. This equivalence between irreducibility of a matrix and the strongly connected property of the graph associated with the matrix is now presented.

The following definitions and a known theorem will be useful. The proof of the theorem is this author's.

Definition 5.6

To any square matrix $A$ of order $n$ there corresponds a directed graph $G(A)$ defined as a set of points $\{P_i : i = 1, 2, \ldots, n\}$ and a set of broken directed lines connecting $P_j$ to $P_k$ whenever $A(j, k) \neq 0$.

Definition 5.7

The directed graph $G(A)$ of a matrix $A$ is said to be strongly connected if it is possible to pass from any one point of the graph to another along the direction lines.

Theorem 5.5

A matrix $A$ is irreducible if and only if its directed graph $G(A)$ is strongly connected.
Proof It is convenient to prove first, the contrapositive, namely, that if \( A \) is reducible then \( G(A) \) is not strongly connected. By Lemma 5.4, if \( A \) is reducible this implies that there exists a permutation matrix \( P \) such that

\[
P AP' = \begin{bmatrix} B & C \\ \alpha_{r+1} & \alpha_{n-r} \\ 0 & D \\ \beta_{n-r} & \beta_{n-1} \end{bmatrix}
\]

We may assume the least favorable situation, namely, that \( B, C, \) and \( D \) are positive matrices. It is sufficient to prove that \( G(PAP') \) is not strongly connected. Consider the points \( P_{r+1} \) and \( P_1 \). Since \( P_{r+1} \) is connected only to points \( P_{r+1}, i \geq 1 \), and any such point \( P_{r+j} \) is in turn connected only to points \( P_{r+1}, i \geq 1 \), it follows that there is no directed path from \( P_{r+1} \) to \( P_1 \), say. Therefore \( G(PAP') \) is not strongly connected. Since \( P \) is a permutation matrix, \( A \) is not strongly connected. Therefore, if the graph of a matrix is strongly connected the matrix is irreducible.

Next, it is proved that if \( G(A) \) is not strongly connected, \( A \) is reducible. Since \( G(A) \) is not strongly connected, there exist two points \( P_i, P_j \) such that no directed path exists from \( P_i \) to \( P_j \). Without loss of generality we may permute the points by interchanging rows and columns of the matrix so that \( i > j \) and \( j = 1 \). Call the resulting matrix \( P_1 A P_j = B \). Clearly \( B(i,1) = 0 \). Now, it must be the case that for every \( k \) such that \( B(i,k) \neq 0, B(k,1) = 0 \). Further, if \( B(k,1) \neq 0 \), then \( B(1,1) = 0 \). The last two implications must hold, for otherwise
it would be possible to connect $P_1$ with $P_1$ via $P_k$ and $P$. For the first $k$ such that $B(i,k) \neq 0$, permute the $k$ and $n$-th rows and $k$-th and $n$-th columns. In general, if the $r$-th $k$ is such that $B(i,k) \neq 0$, then permute the row and column with the $n-r+1$ row and column. The resulting matrix is in reducible form.

**Corollary 5.1**

With the model

$$y_{ijk} = \mu + a_1(i) + a_2(j) + e_{ijk}, \quad B(e_{ijk}) = 0,$$

with $i = 1, 2, \ldots, r_1$; $j = 1, 2, \ldots, r_2$; $k = 0, 1, \ldots, n_{ij}$,

the following conditions are equivalent in that any one implies the other three:

(i) All main effect contrasts are estimable.

(ii) The arrangement of observed cells forms a pairwise connected design.

(iii) The association matrices $N_{12}N_{21}$ or $N_{21}N_{12}$ are irreducible.

(iv) The graph associated with $N_{12}N_{21}$ or $N_{21}N_{12}$ is strongly connected.

The preceding yields various alternative computational methods of determining if the coefficient matrix of a two-way additive arrangement has maximal rank. Of these, when the number of levels of each factor is large, it is especially convenient to verify that the directed graph of $N^1_2$ or $N^2_1$ is strongly connected. Since $N^1_2$ and $N^2_1$ are symmetric, it is possible to graph only the upper triangular part of the matrix using non-directed broken lines. The resulting graph is strongly connected provided there exists a path (non-directed) from every
point to every other point. This is the case if the matrix does not contain a row or column of zeroes above the diagonal.

Suppose now, that the arrangement is not of maximal rank. We need to identify what functions within each factor of classification can be estimated with the given data. The following Corollaries to Theorem 5.2 yield the functions that are estimable with any given arrangement of the observed data.

**Corollary 5.2**

With the model of Theorem 5.2, if the design is not connected, there exists a minimal set of chains \( C_i \) of alternately associated levels of the factors, such that if \( a_m(j) \in C_i \) and \( a_m(j') \in C_j \) then \( a_m(j) \) is not connected to \( a_m(j') \).

**Corollary 5.3**

\( a_m(j) \) and \( a_m(j') \) are elements of the same chain if and only if \( [a_m(j) - a_m(j')] \) is estimable, and in this case, an estimator is given by the same estimator as given in (5.3) based on the connecting chain between the two levels of the factor.

The last two Corollaries to theorem 5.2 and the directed graph associated with the incidence matrix play a central role in the following.

D. The Three Way Additive Classification

This section contains an algorithm that exhibits systematically the functions that can be estimated within each factor of classification when the data available present an arbitrary incidence pattern.
Consider the situation presented by the model

\[(5.4) \quad y_{ijkt} = \mu + a_1(i) + a_2(j) + a_3(k) + e_{ijkt}; \quad E(e_{ijkt}) = 0\]

where \(i = 1, 2, \ldots, r_1; \quad j = 1, 2, \ldots, r_2; \quad k = 1, =, \ldots, r_3;\)

and \(l = 0, 1, \ldots, n_{ijk}\).

An important simplification in the two way model results from the knowledge that estimability of all contrasts of one factor implies and is implied by estimability of all contrasts in the other factor.

In the case under consideration, it is worth considering if estimability of all contrasts in \(a_1(i)\) and all contrasts in \(a_2(j)\) implies the estimability of all contrasts in \(a_3(k)\).

**Theorem 5.6**

In a three way additive classification model, estimability of all linear contrasts within the parameters of any two of the factors implies estimability of all contrasts within the parameters of the third.

**Proof:** Without loss of generality, suppose \(\sum_{i=1}^{r_1} l_{ij} a_1(i)\) and \(\sum_{j=1}^{r_2} m_{ij}\) \(a_2(j)\) are estimable for all sets of real number \(\{l_{ij}\}\) and \(\{m_{ij}\}\) such that

\[\sum_{i=1}^{r_1} l_{ij} = 0 \quad \text{and} \quad \sum_{j=1}^{r_2} m_{ij} = 0.\]

Consider \(\sum_{k=1}^{r_3} n_k a_3(k)\) where \(\sum_{k=1}^{r_3} n_k = 0\).

Since \(\sum_{k=1}^{r_3} n_k a_3(k) = \sum_{k=1}^{r_3} n_k (a_3(k) - a_3(r_3))\), it is necessary only to verify the estimability of each difference \(a_3(k) - a_3(r_3)\). For a fixed \(k\), consider any observation containing \(a_3(k)\) in its expectation,
say, $y_{i_1, j_1, k}$, and an observation containing $a_3(r_3)$ in its expectation, say, $y_{i_2, j_2, r_3}$.

Then $E \left[ y_{i_1, j_1, k} - y_{i_2, j_2, r_3} \right] = (a_1(i_1) - a_1(i_2)) + (a_2(j_1) - a_2(j_2)) + (a_3(k) - a_3(r_3))$.

If $a_1(i_1) \neq a_1(i_2)$, then by the assumptions of the theorem, there exists a vector $\lambda_1$ such that $E(\lambda_1 y) = a_1(i_1) - a_1(i_2)$ where $y$ is the vector of observed values. Similarly, if $a_2(j_1) \neq a_2(j_2)$, then there exists a $\lambda_2$ such that $E(\lambda_2 y) = a_2(j_1) - a_2(j_2)$.

Consider the estimator

$$y_{i_1, j_1, k} - y_{i_2, j_2, r_3} - \lambda_1 y - \lambda_2 y.$$ 

This estimation has $a_3(k) - a_3(r_3)$ as its expectation. Since $k$ was arbitrary it exists for all $k < r_3$ and hence, $\sum_{k=1}^{r_3} n_k a_3(k)$ is estimable for all sets of real numbers $\{n_k\}$ satisfying $\sum_{k=1}^{r_3} n_k = 0$.

Consequently, it is necessary and sufficient to verify the estimability of all contrasts of only two of the three factors. For computational simplicity, the algorithm to be discussed will arrange the factors in ascending order according to the number of levels of each factor, and will verify the estimability of all linear contrasts within each of the first two factors. It is without loss of generality that one can assume $r_1 \leq r_2 \leq r_3$ in model 5.4.
Before proceeding to the development of the algorithm, we make the following general remarks. With the $a_1(i)$, $r_1$ levels of factor $a_1$, there exist at most $r_1 - 1$ independent estimable functions. It is clear that it is necessary and sufficient that $a_1(i) - a_1(r_1)$ be estimable for all $i \leq r_1$, in order that all functions $\sum_{i \neq r_1} a_1(i)$ be estimable. Furthermore, no function can be estimable that is not a linear combination of the $r_1 - 1$ functions in the set $A = \{a_1(i) - a_1(r_1)\}$. It is, however, entirely possible that a linear combination of parametric functions in $A$ is estimable, when the functions in $A$ that make up that combination are not themselves estimable. Thus, for example,

$$a_1(1) - 3a_1(2) + a_1(3) + a_1(4) = (a_1(1) - a_1(2)) + (a_1(3) - a_1(2)) + (a_1(4) - a_1(2))$$

may be estimable even though the components to the right of the equal sign, each of which is a combination of two functions in $A$, are themselves individually non-estimable.

Consider an arbitrary parametric function in the set $A$ represented by $a_1(i) - a_1(r_1)$. Suppose a chain exists between $a_1(i)$ and $a_1(r_1)$ consisting alternately of levels of the factors $a_1$, $a_2$ such that every two adjacent members are associated. Then we have seen in the preceding section how such a chain can be used to define an estimate of $a_1(i) - a_1(r_1)$ with the two factor model. However, there is a difficulty in three factors, important enough to be singled out, namely
that in three factors every estimate does not have a corresponding chain. For example, consider the \( \frac{1}{3} \) replicate of the \( 2^3 \) experiment given by observations with indices corresponding to

<table>
<thead>
<tr>
<th>Sequence number</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(2)</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(3)</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>(4)</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

It is clear that \( \frac{1}{3} [(1) + (2) - (3) - (4)] \) is the only unbiased estimate of \( a_{1}(1) - a_{1}(2) \). This, in the chain notation of the preceding section, would require weighing the estimates defined by the two chains.

\[ C_1 : a_{1}(1), a_{2}(1), a_{1}(2) \]

and

\[ C_2 : a_{1}(1), a_{2}(2), a_{1}(2). \]

For the above reason, an extension of connectedness to the multifactorial additive model in terms of the existence of chains seems futile to the author. Such definitions for connectedness of two levels of two factors are possible, but the best of these seem artificial. These will not be deliberated upon here, and instead the related concepts of irreducibility and graphs will be exploited in what follows.

Consider as an example a realization of the model 5.4, where \( r_1=3, \) \( r_2=4, \) and \( r_3=5 \) are the respective levels of the three factors and the observed cells correspond to indices given by:
<table>
<thead>
<tr>
<th>Cell Sequence Number</th>
<th>i</th>
<th>j</th>
<th>k</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td>2</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>10</td>
<td>3</td>
<td>3</td>
<td>5</td>
</tr>
</tbody>
</table>

The pertinent question is what can be estimated within parameters corresponding to single factors. The following procedure will make use of ordered pairs, in which the elements may at times be levels of factors, and at times, be sequence numbers of cells. To distinguish, we use the notation \((i, j)\) when the elements in the pair are levels of factors, and \(\langle i, j \rangle\) when the elements of the pair are sequence numbers of observed cells. Note that as a general step in the procedure the indices of observed cells \((i, j, k)\) are always arranged lexicographically ignoring repetitions, and then sequentially numbered.

An association matrix \(N_{12}\) can be obtained by starting with a zero matrix of order \(r_1 \times r_2\), then adding 1 to its \((i, j)\) position for each observed cell having the levels of factors 1 and 2 corresponding to \(i\) and \(j\) respectively. For example, for the preceding data some incidence matrices would appear as
Clearly, the incidence matrices are always non-negative, and $N^T_{ij} = N^i_{ji}$. Corresponding to each incidence matrix one can produce a table storing the sequence number of cells corresponding to each positive entry. Thus, in the case of the example, we have

$$\begin{align*}
N_{12} &= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} ;
N_{13} &= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix} ;
N_{23} &= \begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} .
\end{align*}$$

The transpose of the incidence matrix multiplied by the matrix will be denoted $N_{ij} N_{ji} = N^i_j$ and in the case of the example:

$$\begin{align*}
N_{12}^T &= \begin{bmatrix} 4 & 3 & 3 \\ -3 & 2 & 2 \\ -3 & -3 & -3 \end{bmatrix} ;
N_{13}^T &= \begin{bmatrix} 4 & 2 & 2 \\ -3 & 2 & 2 \\ -3 & -2 & 1 \end{bmatrix} ;
N_{23}^T &= \begin{bmatrix} 3 & 2 & 1 & 0 \\ -3 & 2 & 0 \\ -3 & -2 & 1 \\ -3 & -2 & 2 \end{bmatrix} .
\end{align*}$$

The matrices are symmetric and only the upper triangular part need be computed. They are obviously irreducible which is a necessary condition for estimability of all main effect contrasts. However, this is not a sufficient condition as was observed in the preceding section. The tables $M_{ij}$ are multiplied symbolically to produce pairs equal in number to the entries in $N_{ij}^i$. The rule is $M_{ij} \circ M_{ji} = M^i_j$. 
where

\[ M^i_{j}(n,m) = \{(k,l) : k \in M^i_{j}(n,p) \text{ and } l \in M^i_{j}(p,m) \text{ for some } p\} \]

Some resulting tables for the preceding example are

\[
\begin{array}{c|c|c|c|c|c|c|c|c|c}
\hline
\langle 1,1 \rangle & \langle 3,3 \rangle & \langle 1,5 \rangle & \langle 1,8 \rangle \\
\langle 2,2 \rangle & \langle 4,4 \rangle & \langle 2,6 \rangle & \langle 2,9 \rangle \\
\langle 4,7 \rangle & \langle 3,10 \rangle & & & & \\
\hline
\hline
\langle 5,5 \rangle & \langle 6,6 \rangle & \langle 5,8 \rangle & \langle 5,9 \rangle & \langle 6,9 \rangle \\
\langle 7,7 \rangle & & & & & & & & \\
\hline
\hline
\langle 5,5 \rangle & \langle 6,6 \rangle & \langle 6,8 \rangle & \langle 6,9 \rangle & \langle 7,9 \rangle & \langle 8,8 \rangle & \langle 9,9 \rangle & \langle 10,10 \rangle \\
\hline
\end{array}
\]

\[ M_2 = M_{12} \oplus M_{21} = \]

\[
\begin{array}{c|c|c}
\hline
\langle 1,1 \rangle & \langle 3,3 \rangle & \langle 3,3 \rangle \\
\langle 2,2 \rangle & \langle 4,4 \rangle & \langle 4,4 \rangle \\
\hline
\hline
\langle 5,5 \rangle & \langle 6,6 \rangle & \langle 6,6 \rangle \\
\langle 7,7 \rangle & & & & & & & & \\
\hline
\hline
\langle 5,5 \rangle & \langle 6,6 \rangle & \langle 6,6 \rangle \\
\langle 7,7 \rangle & & & & & & & & \\
\hline
\hline
\langle 8,8 \rangle & \langle 9,9 \rangle & \langle 9,9 \rangle \\
\langle 10,10 \rangle & & & & & & & & \\
\hline
\end{array}
\]

\[ M_3 = M_{13} \oplus M_{31} = \]
The resulting tables are symmetric and the lower triangular part would contain the same ordered pairs with the elements of each pair reversed.

It is convenient, now, to digress on the motivation for the above constructions. The off-diagonal entry in the incidence matrix $M_{ij}$ in the $k$ row and $k'$ column, $(k < k')$, is the number of simple differences of two cells which have in their expectation $a_j(k) - a_j(k')$, and do not contain any parameters of $a_i$. Such cell contrasts are implicitly given by any pair in $M_{ij}^1 (k, k')$. For example, $M_{2}^1 (1,3) = 3$, so three simple differences between means of two cells have in their expectations $a_1(1) - a_1(3)$, and do not contain $a_2$ parameters. From $M_{2}^1 (1,3)$ one obtains that $<1,8>$ is one such difference, so that if $\bar{y}(p)$ is the mean of the observations in the cell with sequence number $p$, then

$$E(\bar{y}(1) - \bar{y}(8)) = \left[ a_1(1) - a_1(3) \right] + \left[ \text{simple difference of parameters in } a_2(k) \right].$$

---

<table>
<thead>
<tr>
<th></th>
<th>$\langle 1,1 \rangle$</th>
<th>$\langle 5,2 \rangle$</th>
<th>$\langle 8,3 \rangle$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\langle 2,2 \rangle$</td>
<td>$\langle 6,3 \rangle$</td>
<td>$\langle 10,4 \rangle$</td>
</tr>
<tr>
<td></td>
<td>$\langle 6,6 \rangle$</td>
<td>$\langle 9,7 \rangle$</td>
<td>$\langle 10,10 \rangle$</td>
</tr>
<tr>
<td></td>
<td>$\langle 9,9 \rangle$</td>
<td>$\langle 4,4 \rangle$</td>
<td>$\langle 7,7 \rangle$</td>
</tr>
</tbody>
</table>

$$M_3^2 = M_{32} \odot M_{23} =$$
We proceed to examine some uses of this information.

Consider estimability of \( a_1(1) - a_1(3) \). From \( M_2^1(1,3) \) and \( M_3^1(1,3) \), it is clear that if there exist real constants \( c_i \) such that

\[
E[c_1(\bar{y}(1) - \bar{y}(8)) + c_2(\bar{y}(2) - \bar{y}(9)) + c_3(\bar{y}(3) - \bar{y}(10))] = E[c_4(\bar{y}(3) - \bar{y}(8)) + c_5(\bar{y}(4) - \bar{y}(10))],
\]

then each expectation is equal to \( a_1(1) - a_1(3) \). Every estimator of \( a_1(1) - a_1(3) \) need not be of such form. Graphically, if each cell is represented by a point and each two points corresponding to a pair in \( M_2^1(1,3) \) are connected by a line of one color, while each two points corresponding to a pair in \( M_2^1(1,3) \) are connected by a line of a different color, the resulting graph is

Then \( a_1(1) - a_1(3) \) has an estimator of the above form iff the graph contains a closed loop consisting of an even number of lines alternately chosen from each color. One can obtain similar graphs using the other cells of \( M_2^1 \) and \( M_3^1 \). These are
Examination of the graphs reveals no simple estimable functions of the type $a_1(i) - a_1(i')$. If the graphs $G_1(1,3), G_1(1,2)$, and $G_1(2,3)$ are added, then

A number of loops can be located of the type described above. For example, some are

$L_1: \langle 2 \rangle, \langle 6 \rangle, \langle 8 \rangle, \langle 5 \rangle, \langle 2 \rangle$
Then from the loops, one obtains: that the expectations of

\[ \bar{y}(2) - \bar{y}(6) + \bar{y}(8) - \bar{y}(5), \]
\[ \bar{y}(5) - \bar{y}(8) + \bar{y}(3) - \bar{y}(10) + \bar{y}(4) - \bar{y}(7) + \bar{y}(9) - \bar{y}(2), \]

and \[ \bar{y}(3) - \bar{y}(10) + \bar{y}(4) - \bar{y}(7) + \bar{y}(9) - \bar{y}(6) \]
yield estimable functions in \( a_1(i) \) parameters, estimating

\[ a_1(1) - 2a_2(2) + a_1(3), \quad a_1(1) - a_1(3) \]

and \( 2a_1(1) - 2a_2(2) \). Clearly, the functions are not independent, but do span all estimable functions in \( a_1(i) \) parameters.

Similarly, in factor \( a_2 \) the composite graph is given by:

![Composite graph](image)

So the closed loops defining estimable functions in \( a_2(j) \) parameters are

\[ L_1 : \langle 6 \rangle, \langle 5 \rangle, \langle 2 \rangle, \langle 3 \rangle, \langle 6 \rangle \]
\[ L_2 : \langle 3 \rangle, \langle 9 \rangle, \langle 7 \rangle, \langle 5 \rangle, \langle 2 \rangle, \langle 3 \rangle, \langle 8 \rangle \]
\[ L_3 : \langle 4 \rangle, \langle 3 \rangle, \langle 8 \rangle, \langle 10 \rangle, \langle 4 \rangle. \]
Then the expectations of
\[ \bar{y}(6) - \bar{y}(5) + \bar{y}(2) - \bar{y}(3) \]
\[ \bar{y}(8) - \bar{y}(9) + \bar{y}(7) - \bar{y}(5) + \bar{y}(2) - \bar{y}(3), \]
\[ \bar{y}(4) - \bar{y}(3) + \bar{y}(8) - \bar{y}(10) \]
yield estimable functions in \( a_2(j) \), and these estimate
\[ 2a_2(2) - a_2(1) - a_2(3), \]
\[ a_2(4) - a_2(3), \text{ and} \]
\[ a_2(1) + a_2(4) - 2a_2(3), \text{ respectively.} \]

Then, since these are independent, every contrast in \( a_2(j) \) is estimable.

We have the estimable functions

\[
\begin{bmatrix}
-1 & 2 & -1 & 0 \\
0 & 0 & -1 & 1 \\
1 & 0 & -2 & 1
\end{bmatrix}
\begin{bmatrix}
a_2(1) \\
a_2(2) \\
a_2(3) \\
a_2(4)
\end{bmatrix}
\]

As previously stated, the tables of differences can be viewed in many interesting alternative ways. The following provides a highly efficient algorithm for identifying estimable functions within the parameters of any one factor. Suppose our interest is in the \( a_3(k) \) parameters. \( M_2 \), say, contains all simple differences of cell means whose expectation contain a contrast in \( a_3(k) \) and no \( a_1(j) \) parameters. Now,
Each entry, such as 1,2, defines a simple difference of two cell means $\bar{y}(1) - \bar{y}(2)$, which contains in its expectation a difference in $a_2(j)$ parameters. Such differences can be tabled in an array called $P_1^3$. Then, in the example,

$$
\begin{array}{cccc}
\langle 1,1 \rangle & \langle 1,2 \rangle & \langle 1,3 \rangle & \langle 1,4 \rangle \\
\langle 2,2 \rangle & \langle 2,3 \rangle & \langle 5,7 \rangle & \langle 2,4 \rangle \\
\langle 5,5 \rangle & \langle 5,6 \rangle & \langle 5,7 \rangle & \langle 2,4 \rangle \\
\langle 3,3 \rangle & \langle 6,7 \rangle & \langle 3,4 \rangle & \\
\langle 6,6 \rangle & \langle 8,7 \rangle & \langle 8,9 \rangle & \langle 8,10 \rangle \\
\langle 8,8 \rangle & \langle 9,7 \rangle & \langle 9,9 \rangle & \langle 9,10 \rangle \\
\langle 9,9 \rangle & \langle 9,10 \rangle & \langle 10,10 \rangle & \\
\end{array}
$$

The graphs $G_{1}(k,k')$ of the pair in each cell define estimable functions of $a_2(k)$ parameters as follows. If the graph $G_{1}(k,k')$ contains a closed loop, then $a_2(k) - a_2(k')$ is estimable. Clearly, in the present example no such estimates are found. However, the same criterion can be applied to
\[ G^2(1,2) + G^2(1,3) + G^2(1,4) + G^2(1,5) + G^2(2,3) + G^2(2,4) + \\
+ G^3(2,5) + G^3(3,4) + G^3(3,5) + G^3(4,5). \]

In the example

\[ \frac{r-1}{2} \sum_{k=1}^r \sum_{k'=k+1}^r G^3(k,k'). \]

A few of the closed loops readily found are

- \( L_1 : 1, 2, 1 \)
- \( L_2 : 2, 4, 2 \)
- \( L_3 : 1, 3, 1 \)
- \( L_4 : 1, 4, 1 \)
- \( L_5 : 2, 3, 2 \)

Corresponding functions of cell means from \( \mathbb{R}^2 \) are

\[ \bar{y}(1) - \bar{y}(2) + \bar{y}(6) - \bar{y}(5) \]
\[ \bar{y}(2) - \bar{y}(4) + \bar{y}(7) - \bar{y}(6) \]
\[ \bar{y}(1) - \bar{y}(3) + \bar{y}(10) - \bar{y}(8) \]
\[ \bar{y}(1) - \bar{y}(4) + \bar{y}(7) - \bar{y}(5) \]
\[ \bar{y}(2) - \bar{y}(3) + \bar{y}(10) - \bar{y}(9). \]

The expectations are given by

\[ a_3(1) - 2a_3(2) + a_3(3) \]
\[ a_3(2) - a_3(5) + a_3(4) - a_3(3) \]
\[ a_3(1) - 2a_3(3) + a_3(5) \]
\[ a_3(1) - a_3(5) + a_3(4) - a_3(2) \]
\[ a_3(2) - a_3(3) + a_3(5) - a_3(4) \]

respectively.

These are clearly not independent, but do generate all estimable functions in \( a_3(k) \) parameters.
VI. ACKNOWLEDGEMENTS

I am grateful to my thesis adviser Dr. G. Zyskind for providing much needed direction and for encouraging interest in the topics contained in the thesis.

I sincerely appreciate the assistance of Dr. O. Kempthorne and his contribution through many helpful discussions.

The research presented in this thesis was made possible through the financial assistance of the Statistical Laboratory of Iowa State University. In particular, I am grateful to its Director, Dr. T. A. Bancroft for the opportunity to work and study in the Department of Statistics. Part of this research was also supported by the United States Air Force under Contract Number AF33(615)68-C-1343.

Nothing would have been possible without the moral support of my wife Beverly, who during the last seven years without complaining about a frequently absent husband managed the development of our three children, Michael, Therese and Christopher.
III. BIBLIOGRAPHY


7. Dixon, W. J., ed. BMD: Biomedical computer programs. Health Sciences Computing Facility, Department of Preventive Medicine and Public Health, School of Medicine, University of California, Los Angeles, California. 1964.

8. Elston, R. C. and Bush, N. The hypothesis that can be tested when there are interactions in an analysis of variance model. Biometrics 20: 681-698. 1964.


VIII. APPENDIX

COMPUTER PROGRAM FOR COUNTING THE NUMBER OF STRUCTURES, q = 3

```
1         DIMENSION IRAY(14)
2         COMMON N,IA,II,JPP,JPM1,J,IRAY
3         DO 1 N=4,15
4         C ZERO ACCUMULATORS
5         DO 50 I=1,14
6           50 IRAY(I)=0
7         IA=0
8         NN=N-3
9         DO 2 I=1,NN
10         C CLASS 1 STRUCTURES
11           II=I+1
12           NS0=N-3-I
13           NS1=I
14           DO 3 JP=1,II
15             JPP=JP-1
16             IF(NS0-JPP)3,5,5
17             5 IF(NS1-I+JPP)3,6,6
18             6 CALL SUBA
19             3 CONTINUE
20         C CLASS 2 STRUCTURES
21           NM2=N-2
22           DO 7 J=II,NM2
```
20 \text{JM1=J+1}
21 \text{DO 8 JP=1,JM1}
22 \text{JPM1=JP-1}
23 \text{IF(NSO-JPM1)8,10,10}
24 \text{10 IF(NS1-J+JPM1)8,11,11}
25 \text{11 CALL SUBB}
26 \text{8 CONTINUE}

\text{C CLASS 3 STRUCTURES}
27 \text{DO 9 JP1=1,JM1}
28 \text{JPM1=JP1-1}
29 \text{IF(J-I-JPML)9,12,12}
30 \text{12 IF(J-I-JPML-1)9,13,9}
31 \text{13 CALL SUBC}
32 \text{9 CONTINUE}
33 \text{7 CONTINUE}
34 \text{2 CONTINUE}
35 \text{PRINT 20,N,IRAY,IA}
36 \text{20 FORMAT(1H ,1616)}
37 \text{1 CONTINUE}
38 \text{STOP 0}
39 \text{END}
SUBROUTINE SUBA
DIMENSION IRAY(14)
COMMON N,IA,I,II,JPP,JPM1,J,IRAY
C DEFINES NUMBER OF ELEMENTS IN EACH SET
NS02=N-3-I-JPP
NS1N2=JPP
NS1N1=JPP
NS2= I-JPP
C CLASS 1 STRUCTURES
C CLASS 1.1
DO 100 KP=1,II
KPML=KP-1
IMKP=II-KPML
DO 100 KPP=1,IMKP
KPPML=KPP-1
IF(JPP-KPML)100,2,2
2 IF(NS02-KPML)100,102,102
102 IF(JPP+JPP-KPML)100,103,103
103 IF(NS2-I+KPML+KPPML)100,104,104
104 IA+IA+1
IRAY(1)=IRAY(1)+1
100 CONTINUE
C CLASS 1.2
NML=N-1
DO 111 K=II,NML
KML = K + 1
DO 110 KP = 1, KML
KPM1 = KP - 1
NMKP = K - KPM1 + 1
DO 110 KPP = 1, NMKP
KPPM1 = KPP - 1
IF(NSO2 - KPM1) 110, 112, 112
112 IF(JPF + JPP - KPPM1) 110, 113, 113
113 IF(NS2 - K + KPM1 + KPPM1) 110, 114, 114
114 IA = IA + 1
110 CONTINUE

KT = K - NS2 - NS1N2
C CLASS 1.3
DO 130 KP = 1, KT
KPM1 = KP - 1
IF(NSO2 - KPM1) 130, 132, 132
132 IF(NS1N1 - K + NS2 + NS1N2 + 1 + KPM1) 130, 131, 131
131 IA = IA + 1
130 CONTINUE

C CLASS 1.4
IF(K - NS2 - NS1N1 - NS1S2 - 2) 140, 142, 142
142 IF(NSO2 - K + NS2 + NS1N1 + NS1N2 + 2) 140, 141, 141
141 IA = IA + 1
C

45 \text{IRAY(4)}=\text{IRAY(4)}+1

C

46 140 \text{CONTINUE}

47 111 \text{CONTINUE}

48 \text{RETURN}

49 \text{END}
SUBROUTINE SUBB

DIMENSION IRAY(14)

COMMON N,IA,I,II,JPP,JPM1,J,IRAY

C DEFINES THE NUMBER OF ELEMENTS IN EACH SET

NS02 = N-3-I-JPM1
NS1N2 = I-J+JPM1
NS1N1 = JPM1
NS2 = J-JPM1
IF(NS1N2) 2,1,1
2 NS1N2 = 0
1 IF(J-I-JPM1) 3,3,4
4 NS2 = NS2-1

C CLASS 2 STRUCTURES

C CLASS 2.1

JML = J+1
DO 100 KP = 1, JML
KPM1 = KP-1
JMKPL = J-KP+2
DO 100 KP2 = 1, JMKPL
KP2ML = KP2-1
JMK12 = J-KP1-KP2+2
DO 100 KP3 = 1, JMK12
KP3ML = KP3-1
IF(KPML-NS02) 10,10,100
22 10 IF(KP2M1-NS1N2)12,12,100
23 12 IF(KP3M1-NS1N1)13,13,100
24 13 IF(J-KPM1-KP2M1-KP3M1-NS2)14,14,100
25 14 IF(KPML-JPML+KP3ML)502,502,100
26 502 IA=IA+1
27 IRAY(5)=IRAY(5)+1
28 100 CONTINUE

C CLASS 2.2

29 NS1N2=I-J+JPML
30 JMI1= J-1
31 DO 110 KP=1,JMI1
32 KPML=KP-1
33 IF(KPML-NS02)15,15,110
34 15 IF(J-I-1-KPM1-NS1N1)16,16,110
35 16 IA=IA+1
36 IRAY(6)=IRAY(6)+1
37 110 CONTINUE

C CLASS 2.3

38 NML=N-1
39 JML=J+1
40 DO 130 K=JML,NML
41 KMI=K-J
42 DO 131 KP=1,KMI
43 KPML=KP-1
44 IF(KPML-NS02)17,17,131
45 IF(K-KPML-1-NS2-NS1N1-NS1N2)18,18,131
46 IA=IA+1
47 IRAY(7)=IRAY(7)+1
48 CONTINUE

C CLASS 2.4

49 KMJ=K-1
50 DO 132 KP=1,KMJ
51 KPM1(KP-1
52 IF(KPML-NSO2)19,19,132
53 IF(K-1-NS2-NS1N1-NS1N2-KPML)20,20,132
54 IRAY(8)=IRAY(8)+1
55 CONTINUE

C CLASS 2.5

57 KML=K+1
58 DO 133 KP=1,KML
59 KPML(KP-1
60 KMP=K-KP+2
61 DO 133 KP2=1,KMP
62 KP2ML(KP2-1
63 KMPMP=K-KP-KP2+3
64 DO 133 KP3=1,KMPMP
65 KP3ML(KP3-1
66 IF(KPML-NSO2)21,21,133
67 IF(KP2ML-NS1N2)22,22,133
68 IF(KP3ML-NS1N1)23,23,133
212

69  23 IF(K=KPM1-KP2M1-KPSM1-NS2)24,24,133
70   24 IA=IA+1
71     IRAY(9)=IRAY(9)+1
72  133 CONTINUE
73  130 CONTINUE

C CLASS 2.6

74   IA=IA+N-I-JPM1-2
75   IRAY(10)=IRAY(10)+N-I-JPM1-2
76   RETURN
77   END