MINIMAL MULTIDIMENSIONAL DESIGNS

by

John T. Sennetti

Thesis submitted to the Graduate Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
in
Statistics

APPROVED:

Dr. Klaus Hinkelmann, Chairman

Dr. Boyd Harshbarger
Department of Statistics

Dr. Clyde Kramer
Department of Statistics

Dr. Ernest Jennelle
Department of Civil Engineering

Dr. Raymond Myers
Department of Statistics

July, 1972

Blacksburg, Virginia
ACKNOWLEDGEMENTS

I would like to express my appreciation to Dr. Klaus Hinkelmann for his patience, timely advice, and scholarly guidance in this research. Appreciation is also expressed to Dr. Boyd Harshbarger whose dedication to the success of his graduate students will long be remembered, to Dr. Clyde Kramer whose suggestions helped clarify the analysis and organize its presentation, to Dr. Raymond Myers for his contributions and interest and to Dr. Ernest Jennelle for his moral support and encouragement. Sincere gratitude is given to the National Institute of Environmental Health Sciences which supported this investigation under NIH Training Grant, numbers 7T01 ES000 33-08 and 7T01 ES000 33-09.

Finally, I wish to thank my wife, , for her patience, endurance, and encouragement during the preparation of this thesis. Heartfelt appreciation is due to , who undertook the arduous task of typing the rough draft in its various stages and the manuscript in its final form. It may be truly said that this work could not have been accomplished without her untiring support. Therefore, this thesis is dedicated to her.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>1.1 Minimal Designs</td>
<td>2</td>
</tr>
<tr>
<td>1.2 Augmenting Designs</td>
<td>5</td>
</tr>
<tr>
<td>1.3 Optimal Designs</td>
<td>7</td>
</tr>
<tr>
<td>1.4 Literature Review</td>
<td>7</td>
</tr>
<tr>
<td>1.5 Comments on The Available Literature</td>
<td>10</td>
</tr>
</tbody>
</table>

| II. THE ANALYSIS OF DESIGNS WHERE NO INTERACTION IS ASSUMED | 12 |
| 2.1 Introduction to the Analysis Of Two-Factor Designs | 13 |
| 2.2 The Analysis of m-Factor Designs | 20 |
| 2.3 A General Form For Calculating C_a, The Coefficient Matrix In The Reduced Normal Equation For P_a | 23 |
| 2.4 Calculating Generalized Inverses | 27 |

| III. THE ANALYSIS OF DESIGNS WHERE SOME TWO-FACTOR INTERACTIONS ARE ASSUMED | 30 |
| 3.1 The Two Factor Design With Interaction | 32 |
| 3.2 The m-Factor Design With One Interaction | 37 |
| 3.3 The m-Factor Design With Two Interactions | 41 |
| 3.4 The m-Factor Design With Three Interactions | 48 |

| IV. CONNECTING EXPERIMENTAL DESIGNS | 55 |
| 4.1 A Literature Review and Summary of Results | 55 |
| 4.2 Augmenting Designs With Assumed First Order Interactions | 56 |

| V. THE OPTIMALITY OF MINIMAL AND MINIMAL AUGMENTED MULTIDIMENSIONAL DESIGNS | 72 |
| 5.1 Introduction | 72 |
| 5.2 Summary of Chapter V | 73 |
| 5.3 MMD's and MAMD's | 73 |
| 5.4 A Review Of Some Main Effect Optimality Criteria | 75 |
| 5.5 Finding Optimal MAMD's | 77 |
| 5.6 Reasons For Finding MMD's and MAMD's That Are \( \gamma \)-Optimal For Only One Set, \( S_\alpha \) | 84 |
| 5.7 Procedure I | 86 |
| 5.8 Procedure II | 97 |
| 5.9 Examples of Minimal Designs With Two-Factor Interactions | 112 |
TABLE OF CONTENTS (Continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI. PROPOSED EXTENSIONS</td>
<td>126</td>
</tr>
<tr>
<td>VII. SUMMARY</td>
<td>129</td>
</tr>
<tr>
<td>POSTSCRIPT</td>
<td>132</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>133</td>
</tr>
<tr>
<td>APPENDIX I</td>
<td>137</td>
</tr>
<tr>
<td>APPENDIX II</td>
<td>148</td>
</tr>
<tr>
<td>VITA</td>
<td>153</td>
</tr>
</tbody>
</table>


LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. A Main Effect Plan For A 3x2x4x6 Factorial In 12 Assemblies: $T_1$</td>
<td>104</td>
</tr>
<tr>
<td>II. The Construction of $T_4$</td>
<td>106</td>
</tr>
<tr>
<td>III. Assemblies That Connect $S_0$ and $S_3$</td>
<td>108</td>
</tr>
<tr>
<td>IV. The Set $\psi$ for a 3x2x6 Factorial</td>
<td>109</td>
</tr>
<tr>
<td>V. The Sum, $\frac{1}{\lambda_i}$, For Members of $\psi$, Example 5.4.</td>
<td>111</td>
</tr>
<tr>
<td>VI. A $\psi$-Optimal Design For Case (1)</td>
<td>113</td>
</tr>
<tr>
<td>VII. A Set $\psi$ For Case (2)</td>
<td>115</td>
</tr>
<tr>
<td>VIII. The Sum, $\frac{1}{\lambda_i}$ For Members of $\psi$, Case (2)</td>
<td>117</td>
</tr>
<tr>
<td>IX. The Design $T_3$ of Case (2)</td>
<td>118</td>
</tr>
<tr>
<td>X. 11 Assemblies Of the Required 15, Case (3)</td>
<td>121</td>
</tr>
<tr>
<td>XI. One Design Of 15 Assemblies For Case (3)</td>
<td>122</td>
</tr>
<tr>
<td>XII. One Design Of 19 Assemblies For Case (4)</td>
<td>124</td>
</tr>
<tr>
<td>XIII. Sales Data From Preliminary Test On Frozen Fruit Juices</td>
<td>139</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

An experimental design may be considered an arrangement of the degrees (levels) of identified forces (factors) in a plan of investigation. The purpose of this arrangement is to control or measure factorial effects on experimental units to which factor level combinations are applied. Multidimensional designs (MD's) are multifactor fixed effect experimental designs. Those MD's where not all factor level combinations occur (which we will call incomplete designs) shall be the subject of our investigation. An example of such a design is a Latin square where rows, columns, and treatments are considered fixed effect factors, each at k levels. Other examples include Graeco-latin squares, lattice designs, balanced and partially balanced incomplete block designs, F-squares, and multidimensional partially balanced designs.

The above designs are used when no interaction is assumed between any factors such as treatments, blocks, rows, columns, etc. This is not always a valid assumption. In those cases where this assumption may not hold, it is natural to investigate designs which allow for the presence of certain two-factor interactions. Some designs which allow for this possibility are given by Potthoff (1958, 1962), Causey (1967, 1968), and Anderson (1968). These latter designs are special cases of fractional factorials where some interactions are assumed negligible. Therefore MD's are fairly common. Researchers in industry, education, and almost all scientific fields have used these plans of investigation. (Potthoff, 1963; Bose and Srivastava, 1964a)
1.1 Minimal Designs

One problem that arises in such plans of investigation is that the number of factors, and thus the number of factor level combinations, may become too large to carry out. In an attempt to economize on experimental units, experimenters use experience or prior knowledge concerning the factors involved and hence choose designs to estimate only certain effects considered important. An experimenter may assume two-factor and higher order interactions to be negligible. Orthogonal main effect plans, for example, permit the estimation of only the main effects. Assuming no interactions present, one may obtain these estimates with a desired property, orthogonality. Also, some economy of experimental units is achieved in these designs. However, with the same number of experimental units or even less, it may be possible to allow for two-factor interactions, but give up orthogonality. Having information on two-factor interactions may be considered more important than having orthogonal main effect estimates. Perhaps in using the experimenter's knowledge about the assumed effects it may be possible to emphasize economy entirely. Presently there is no literature on experimental designs that (1) permit the estimation of all assumed main effects and two-factor interactions and (2) use as few assemblies (factor level combinations) as possible. The MD's available require more assemblies than absolutely necessary. These "extra" assemblies provide such properties as orthogonality and balance. By ignoring these properties, we may require fewer assemblies.
Consider, as a point of illustration, an example given by Potthoff (1963). The example concerns a pricing experiment on certain types (A, B, C) of frozen fruit juice. This experiment will be a preliminary test which will provide information for future tests. The experimenters wish to determine the effect of three different prices (1, 2, 3) on the sales of each of three different types of fruit juices. The experiment is to take place on the days, Thursday, Friday, and Saturday (I, II, III), when store sales are the greatest. Thus the three factors in the experiment are price, type of fruit juice, and day.

A complete factorial experiment would require that each type of juice be sold every day at all three prices, simultaneously. Potthoff (1963) says,

This would be out of the question: the same type of juice could not be offered for sale in the store at more than one price at any one time. Hence it is necessary to consider using an incomplete design which would specify the selling price of each type of juice for each day of the experiment.

Potthoff (1963) in this example later considers one type of possible two-factor interaction, a price-juice interaction. We shall, for the purpose of illustration, assume there is no such interaction. A price-juice interaction would mean that lower prices would increase the sales of some juice more than of others. Also, the experimenters assume no price-day interaction:

The socio-economic characteristics of the shoppers are not different depending on the day of the week and/or the weather, and a price elasticity of demand is not assumed to be different for the socio-economic groups of consumers.
As for a juice-day interaction, the experimenters are willing to assume that the choice of juice would not be influenced by the day of the week, weather, or socio-economic factors. From this example and these assumptions, we will now begin to illustrate a certain type of design.

Assuming no interaction occurs between or among types of fruit juices, prices or days, two degrees of freedom (d.f.) are required for each factor in order to estimate the differences in juice types, pricing levels, and days. Therefore a total of seven particular periods involving only the days Thursday, Friday, and Saturday, and different price-juice level combinations, are needed to estimate the desired effects. An orthogonal main effect plan may be considered here: a $1/3$ replicate of a $3^3$ factorial using the factor level combinations from a $3 \times 3$ Latin square. However, this design requires nine level combinations (assemblies) and only seven are absolutely necessary. These seven could be the level combinations $(1AI)$, $(2BII)$, $(3CII)$, $(1BI)$, $(2CIII)$, $(3CIII)$, and $(1AII)$. More shall be said on this type of design in Chapter V.

Since careful selection and monitoring of particular test periods will incur considerable expense, it is desired that as few periods as possible be used, especially since the study will be a preliminary test. Further, it may be that only very few periods are available for such a test.

For the moment let us ignore experimental error. A design utilizing only seven periods would have to be non-orthogonal and could not
have the usual analysis of a Latin square, but would at least be preferred to not experimenting at all, if this were the choice. An estimate of error then, could perhaps come from prior information (see Cochran and Cox, 1957, p.262) or some other source. Note that with only eight test periods available a design using seven periods would still apply. If 9 periods were still available, it still is feasible then to perform a particular set of seven necessary to estimate desired effects. Given the necessary seven assemblies, any more assemblies would then provide d.f. for experimental error.

This fruit juice study illustrates the type of problem one may encounter in constructing and analyzing a design with a minimum of factor level combinations: a minimal design. The real usefulness of minimal designs is made more apparent when we consider many factors each with many levels. For these cases even "economic" designs (Bose and Srivastava, 1964a) may require more assemblies than necessary. Our interest will be directed toward minimal designs of \( m \) factors \( F_i \), with \( n_i \) levels, \( i = 1, \ldots, m \), that permit not only the estimation of main effects, but certain assumed interactions as well. Chapter V will present a class of such designs with examples that assume two factor interaction models.

1.2 Augmenting Designs

Parallel to the need for a minimal design is the need for a minimum set of new factor level combinations added to an established design for the purpose of comparing effects not estimable or not assumed present in
the original design. For example, in many factorial main effect plans, the question is occasionally raised on the possible presence of two-factor interactions. The orthogonal main effect plans of Addelman and Kempthorne (1961) enable an experimenter to estimate the differences among levels of each factor using some degree of economy in level combinations. Therefore a procedure for adding factor level combinations to such a main effect plan to allow for interaction would be highly desirable.

This is particularly true, if due to the limitations of experimental units, the usual factorial fractions (incomplete factorials) do not apply. The advantage here is to make use of the existing main effect design rather than to construct new factorial fractions which may require more assemblies than really necessary. Also, since incomplete designs are chosen for their characteristic economy of experimental units, it may seem especially desirable to the experimenter if the factor level combinations to be added are minimum in number; i.e., minimal. Such an augmented design would provide information on suspected interaction and future experiments may then be performed in light of this information.

Let us consider the study of fruit juices mentioned earlier. Assume that the nine assemblies necessary for the orthogonal 1/3 replicate are available. The questions may be raised on what assemblies (and thus periods) are to be added to the original nine in order to allow for the possible effects of a price-juice interaction. Since test periods may be at a premium, the fewest periods necessary may be
preferred. If a preliminary test is made and no significant interaction noted, future studies may require fewer periods than used by the augmented design.

For example, let the nine assemblies of a 1/3 replicate be represented by (lAI), (2BI), (LCII), (2AII), (3BII), (LBIII), (2CIII), and (3AIII). The pair of assemblies (lBII) and (lBI) could be used to augment the original nine for the presence of a price-juice interaction. More shall be said of this in Chapter IV.

1.3 Optimal Designs

The very ability to construct a minimal design or a procedure to augment designs with the minimum number of assemblies may be in itself an optimality criterion. Certainly these designs may not have the orthogonal properties of common MD's. One must pay a price for using less information or fewer experimental units. However, given this constrain of limited assemblies, we may choose from a set of available minimal designs a design that is optimal according to common optimality criteria. For example, from a set of augmented designs we may determine which augmentation procedure enables one to estimate certain desired effects with the least average variance. Other common optimality criteria shall be discussed in Chapter V. Methods of constructing "optimum" sets from which one chooses an optimal design shall be presented also.

1.4 Literature Review

Before pursuing our investigation any further, let us provide some foundation for our inquiry. As mentioned, no previous work is given in
the literature where development is directed specifically toward minimal designs. However, we may present some related works in the analysis and construction of some MD's where main effects and certain interactions are of interest. We shall limit our review to this area of interest rather than present a review of extensive development in experimental designs that includes investigation by Fisher (1935,), Yates (1936), Youden (1937), Cornish (1938) and many more. Kempthorne (1952), Scheffe (1959), and Hinkelmann (1968a) cover most of this work to date. After discussing useful research in the analysis and construction of MD's, we will present previous work in augmenting MD's.

Potthoff (1958) first used the term "multidimensional" to describe multifactor fixed effect designs. His investigations are in the construction and analysis of designs that estimate main effects and some interactions, and involve some economy of experimental units.

Potthoff (1958, 1962a, 1962b, 1963) constructed his design classes from Kronecker products of balanced and partially balanced incomplete block designs. This enabled Potthoff to develop from the normal equations a reduced form for one factor similar to the reduced from one receives for treatments in the analysis of incomplete block designs. However, Potthoff's designs are not minimal and are actually special cases of a more general class: multidimensional partially balanced designs (MPBD's) first described by Bose and Srivastava (1964b) and later by Anderson (1968). Bose and Srivastava present an algorithm for solving the full normal equations which result from such designs. This analysis, although more general than Potthoff's is considered tedious
and unappealing (Hoke, 1971). Further, MPBD's are not minimal designs. However, in the construction of some MPBD's, Anderson (1968) indicated some helpful theorems on the concept of connected designs which shall later prove useful in constructing and augmenting MD's.

Interest toward minimal designs is currently increasing. Daniel (1971) presents a minimal main effect $2^3$ factorial plan. Minimum designs were the subjects of discussion by Federer, Hedayat, and Raktoe (1972), and more work in the area of minimal designs is expected to appear.

The concept of augmenting designs is not new. Methods of adding factorial fractions to established designs in order to estimate effects not estimable or not assumed present have been given by Daniel (1962), Addelman (1963), and John (1966). These methods involve constructing assemblies in order to break the alias chains of two factor interactions, so that certain interaction effects may be estimated. However, the direction of this work is toward developing balanced fractions, rather than fractions involving the minimum number of assemblies.

Augmenting block designs with more treatments (and thus factor level combinations of treatments and blocks) is discussed by Federer (1956, 1961). Federer develops a procedure for adding new treatments (more levels of one factor) to a design and thus the nature of augmenting is slightly different. However, the addition of these new treatments is achieved with the minimum of experimental units. Also if one considers the set of treatments (old and new) as levels of one factor, Federer's work for fixed effects then may be considered a special
1.5 Comments On The Available Literature

In reviewing the current literature on experimental designs, one observation can not be overlooked. It is noted by J. T. Tocher in the critical review following Kiefer (1959) that the construction of many designs involves such mathematical expertise, that experimenters often view such construction more as a mathematical game, rather than as a practical plan of investigation. Consequently, it is not uncommon for a researcher, trained in agriculture for example, not to be familiar with a Galois field, hypersphere, or partially balanced array. (For example, see Vadja, 1967).

Our research has been motivated, therefore, by strong practical considerations. Due to unique experimental constraints on factor level combinations or on the number of experimental units, common MD's often do not apply. Hopefully, since the designs about to be discussed require fewer experimental constraints, an experimenter will be able to fit a design to the problem at hand, rather than change the problem to fit available designs.

In accordance with the experimenter's viewpoint, we shall try to establish designs which are also easy to construct and analyze.

Attempts to simplify analysis are given in Chapters II and III. In these Chapters we shall present a general form for the reduced normal equations. This method generalizes Potthoff's approach and is applicable to all non-orthogonal, incomplete, fixed effect designs,
where two-factor interactions may or may not be assumed present. By reducing the equations to be solved, the analysis is made simpler. In place of the analysis given by Bose and Srivastava (1964b) our analysis is recommended. Also, our approach to the analysis serves as a guide to the construction of minimal designs given in Chapter V.

By extending the work of Anderson (1968) to include interaction models, and by applying these results, we present a new approach to the construction of experimental designs in Chapter IV. With this new approach, an experimenter can construct the needed factor level combinations directly, rather than first investigate such mathematical concepts as vector spaces or investigate designs which may require changing the experiment. In summary then, we hope to present statistical designs and methods that will be easily grasped by any experimenter.

A good factorial fraction is said to have (1) economy, (2) little correlation between desired estimates, (3) fairly uniform variance on these estimates, and (4) an easy method of application and analysis (Bose and Srivastava 1964a). Fractions with all four properties are not at all common. In order to achieve any one property, other properties may have to be sacrificed.

For minimal designs, (2) and (3) may be difficult to achieve. Therefore, in our work we shall concentrate on properties (1) and (4), but shall not ignore (2) and (3) completely. It is with this philosophy that we construct minimal designs and augment designs with a minimum of experimental units.
CHAPTER II

THE ANALYSIS OF DESIGNS WHERE NO INTERACTION IS ASSUMED

Designs which estimate only main effects and use the minimum number of experimental units, may not be orthogonal. If this is true, then an analysis more general than that of orthogonal designs is necessary. In this chapter we shall develop a special approach to the analysis of non-orthogonal m factor designs where no interaction is assumed present.

The purpose of this special approach is to simplify the analysis of the design. This method also provides a device for the construction of the minimal designs discussed in Chapter V. However, this approach shall be similar in part to that of Potthoff (1958).

Potthoff's approach follows the usual procedure for finding treatment estimates adjusted for blocks in incomplete block designs (for example, see Hinkelmann, 1968a). Using this approach, Potthoff finds the sum of squares due to one factor adjusted for m-1 factors for an m factor design. This sum of squares results from a reduced form of the full normal equations (i.e. solving the reduced normal equations and multiplying these solutions by the right hand side of the reduced normal equations). Potthoff developed his analysis for only partially balanced designs. We shall not restrict ourselves to partially balanced designs, but provide an analysis for all incomplete designs and thereby generalize Potthoff's results.

This chapter shall discuss the following: Section 2.1 presents
the two-factor case in order to better understand the analysis of m factors; the usefulness of generalizing this analysis is given in section 2.2; and section 2.3 presents a general form for the reduced normal equations for m factor fixed effect models; and methods for computing this form are given in section 2.4.

2.1 An Introduction To The Analysis Of Two-Factor Designs

We shall first establish notation which shall serve us later for m factors. Assume we have a two-factor design such that a typical observation is represented by \( y_{ij\omega} \), where \( i = 1, \ldots, n_1 \), \( j = 1, \ldots, n_2 \), and \( \omega = 0, 1, \ldots, h_{ij} \). Here \( h_{ij} \) is an element of the incidence matrix \( H = (h_{ij}) \), and represents the number of times a combination \((i,j)\) is replicated (or the number of times factor one at level \( i \) appears with factor two at level \( j \)). Note that if \( h_{ij} = 0 \), \( y_{ij\omega} \) does not appear in the set of observations. Let us assume error terms are normally, independently, homoscedastically distributed such that \( y_{ij\omega} \sim N(\mu, \sigma^2) \); for purposes of estimation the normality assumption can be dropped.

For the two factor no interaction model, \( E(y_{ij\omega}) \) is written as

\[
E(y_{ij\omega}) = \mu + f^i_1 + f^j_2 ,
\]

where \( \mu \) is the general mean and \( f^i_1 \) and \( f^j_2 \) are the \( i \)-th and \( j \)-th level effects of the first and second factor respectively. We may rewrite (2.1.1) as
where, for example, \( p_{i}^{1} f_{1} \) and \( p_{j}^{2} f_{2} + \mu \). Also, for convenience, we shall use standard matrix and vector notation. Therefore, let

\[
p_{1}' = (p_{1}^{1}, \ldots, p_{n_{1}}^{1}) \quad \text{and} \quad p_{2}' = (p_{1}^{2}, \ldots, p_{n_{2}}^{2})
\]

be the factor parameters in the model, and \( X = [X_{1} : X_{2}] \) be the design matrix with respective design matrices, \( X_{1} \) of order \( h \times n_{1} \) and \( X_{2} \) of order \( h \times n_{2} \), to be associated with \( p_{1}' \) and \( p_{2}' \). In defining \( X_{1} \) and \( X_{2} \) in this manner, \( H_{12} \) may then be expressed as

\[
H_{12} = X_{1}' X_{2} = (h_{ij}).
\]

Let \( h = \Sigma h_{ij} \) and \( y \) be the \( h \times 1 \) observation vector. Then

\[
E(y) = X_{1}' p_{1} + X_{2}' p_{2}
\]

where

\[
y' = (y_{11}, y_{12}, \ldots, y_{1h_{11}}, y_{12}, \ldots, y_{1h_{12}}, \ldots, y_{n_{1}n_{2}}, y_{21}, \ldots, y_{2n_{2}})
\]

Further, let us define factor totals which will be used later:

\[
X_{1}' y = Y_{1} \quad \text{and} \quad X_{2}' y = Y_{2}
\]

with

\[
Y_{1}' = (Y_{11}, Y_{12}, \ldots, Y_{1n_{1}}),
\]

\[
Y_{2}' = (Y_{21}, Y_{22}, \ldots, Y_{2n_{2}})
\]
and

\[
Y_{11} = \sum_{j} \sum_{\omega} y_{1j\omega}; \ldots; Y_{1n_1} = \sum_{j} \sum_{\omega} y_{n_1j\omega},
\]

\[
Y_{21} = \sum_{i} \sum_{\omega} y_{1i\omega}; \ldots; Y_{2n_2} = \sum_{i} \sum_{\omega} y_{in_2\omega}.
\]

Let \( \Delta \) be a linear unbiased estimator of some function of the first factor effects only. \( E[\Delta] \), then, must not contain any second factor effects. Since \( \Delta \) is a linear estimate, it must represent a linear combination of the observations, as for example

\[
\Delta = q_1' Y_{11} + q_2' Y_{21} + \sum_{i} \sum_{\omega} y_{1i\omega}
\]

or

\[
\Delta = \sum_{i} q_{1i} Y_{11} + \sum_{j} q_{2j} Y_{21} + \sum_{i} \sum_{\omega} y_{1i\omega}
\]

where the \( q_{1i} \) are vectors of real numbers and

\[
q_1' = (q_{11}, q_{12}, \ldots, q_{1n_1})
\]

\[
q_2' = (q_{21}, q_{22}, \ldots, q_{2n_2})
\]

Using first (2.1.6) and then (2.1.5) and (2.1.3), we can express \( E(\Delta) \) as:

\[
E(\Delta) = E(q_1' Y_{11} + q_2' Y_{21})
\]

\[
= E([q_1' X_{11} + q_2' X_{21}] Y)
\]
\[ E(\Delta) = (q'_1 X'_1 + q'_2 X'_2)(X_{12} p + X_{22} p) \]

\[ = (q'_1 X'_1 + q'_2 X'_2) X_{12} p + (q'_1 X'_1 + q'_2 X'_2) X_{22} p \]

We observe here that \( E(\Delta) \) will be free of factor two if and only if

\[(2.1.10) \quad q'_1 X'_1 X_{12} + q'_2 X'_2 X_{22} = 0' \]

If we let \( H' = X'_1 X' \), \( H = X'_1 X \), and \( H'' = X'_1 X \) we may write

\[(2.1.10) \text{ as follows} \]

\[(2.1.11) \quad q'_1 H_{12} + q'_2 H_{22} = 0 \]

We should note that the above notation is found in the normal equation as

\[(2.1.12) \]

\[
\begin{bmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{bmatrix}
\begin{bmatrix}
p_1 \\
p_2
\end{bmatrix}
= 
\begin{bmatrix}
y_1 \\
y_2
\end{bmatrix}
\]

Note that \( H_{11} = \text{diag}(h_{1.}) \) and \( H_{22} = \text{diag}(h_{.2}) \), with

\[
h_{1.} = \sum_{j=1}^{n_2} h_{1j} \quad \text{and} \quad h_{.2} = \sum_{i=1}^{n_1} h_{ij}.
\]

For the purpose of analysis, Potthoff (1958) requires \( h_{1.} \), the occurrence of the i-th level of factor one, to be equal to \( \frac{h}{n_1} \), and similarly \( h_{.2} = \frac{h}{n_2} \); whereas Anderson (1968) requires \( h_{1.} = \alpha_1 \), be a constant for all
levels of factor one, and similarly $h_{j} = \alpha_{2}$, a constant for factor two.

Note that we require no such restriction here.

From (2.1.11) we get

\[ q'_{2} = -q'_{1} H_{12} H_{22}^{-1} \]

Using (2.1.13) in (2.1.9) we then may express $E(\Delta)$ as a function of $p_{1}$ only:

\[ E(\Delta) = (q'_{1} X'_{1} - q'_{1} H_{12} H_{22}^{-1} X'_{2}) X_{1} p_{1} \]

\[ = q'_{1} (H_{11} - H_{12} H_{22}^{-1} H_{21}) p_{1} \]

\[ = q'_{1} C_{1} p_{1} \]

where the term involving $p_{2}$ is zero, and $C$ is defined:

\[ C_{1} = H_{11} - H_{12} H_{22}^{-1} H_{21} \]

The $(i,I)$ element of $C$ is given by

\[ c_{iI} = h_{i} \delta_{iI} - \sum_{j} \frac{h_{ij}^{2} h_{ij}}{h_{ij}} \quad (i,I = 1,2,...,n_{1}) \]

where $\delta_{iI}$ is the Kronecker delta. If we substitute (2.1.13) into (2.1.6) we get

\[ \Lambda = q'_{1} (Y'_{1} - H_{12} H_{22}^{-1} Y_{2}) \]

\[ = q'_{1} Q_{1} \]
where

\[ Q' = (Q_{11}, Q_{12}, \ldots, Q_{ln}) = (Y_{1} - H_{12} H_{12}^{-1} Y_{2})' \]

Taking (2.1.14) and (2.1.15) we find

\[ q' E(Q) = q' C p \]

or since the above is true for any \( q_1' \),

\[ E(Q) = C p \]

so that for a solution \( p_1 \) to the above we have

\[ Q_1 = C p_1 \]

A solution to (2.1.16) is a familiar result in least squares theory, (see Hinkelmann, 1968a) \( \hat{p}_{1} = C^{-} Q_{1} \), where \( C^{-} \) is a matrix such that

\[ C C^{-} C = C \]

\( C^{-} \) is called a generalized inverse (g-inverse) of \( C_{1} \). More shall be said of g-inverses in section 2.4. Let \( r(C) = n - r \) be the rank of \( C_{1} \). Note that \( n - r < n - 1 \) since

\[ \sum_{i} c_{i1} = \sum_{i} c_{i2} = 0 \]

If \( r = 1 \), the design is said to be connected and every contrast
in \( p \) is then estimable. If \( L' p \) is a contrast, its best linear unbiased estimate is \( L' \hat{p} \), and the variance of \( L' \hat{p} \) is \( L' C^{-1} L \sigma^2 \).

Finding a matrix \( C^{-1} \) is therefore quite important for analysis, as we see later.

In a manner similar to Potthoff (1958), we develop the analysis of variance of our two-factor model, and then extend this development to \( m \) factor models. The development is based on the conditional error principle which is well known and can be found in Scheffe (1959, Chapter 2).

Under the hypothesis that all first factor effects are equal to \( g \), say, (2.1.3) becomes

\[
E(y) = X p^* \quad \text{where} \quad p^* = p^0 + gJ_{n1} \]

and \( J_{n1} \) is a \( n \times 1 \) unity vector.

The conditional error sum of squares (s.s.) under this hypothesis shall be denoted by \( ss_{ce} \) and is

\[
ss_{ce} = y'y - Y^' H^{-1} y^2 \]

From the reduced normal equations (r.n.e.) in (2.1.16), \( \hat{p} \) is found
and the s.s. for factor one is

\[ ss_{p_1} = \hat{p}' Q_{1} \]

The error s.s. under model (2.1.3) is then

\[ ss_e = ss_{ce} - ss_{p_1} = y'y - y' H^{-1} y - \hat{p}' Q_{1} \]

Therefore in summary,

<table>
<thead>
<tr>
<th>Source</th>
<th>d.f.</th>
<th>s.s</th>
<th>E(MS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factor One</td>
<td>$n - r$</td>
<td>$\hat{p}' Q_{1}$</td>
<td>$\frac{p' C_p + \sigma}{n - r}$</td>
</tr>
<tr>
<td>Factor Two</td>
<td>$2$</td>
<td>$y' H^{-1} y^{2}$</td>
<td></td>
</tr>
<tr>
<td>Error</td>
<td>$h - n + r$</td>
<td>$y'y - y'H^{-1} y - \hat{p}' Q_{1}$</td>
<td>$\sigma$</td>
</tr>
<tr>
<td>Total</td>
<td>$h$</td>
<td>$y'y$</td>
<td></td>
</tr>
</tbody>
</table>

2.2 The Analysis of $m$-Factor Designs

Before we extend the results of section 2.1, let us review what has been presented. The results of section 2.2 can be found essentially in Hinkelmann (1968a). This analysis represents no contribution to what is already known, but this development is given to provide a basis for
extension to m-factor models.

Let us now extend our results. Assume an m-factor design with a typical observation given by $y_{ij\cdots z\omega}$, where $i=1,\ldots,n_1$, $j=1,\ldots,n_2$, $z=1,\ldots,n_m$, and $\omega = 0,1,\ldots,h_{ij\cdots z\omega}$. Again $h_{ij\cdots z\omega}$ represents the number of times the combination of factor levels $(i,j,\ldots,z)$ appears, and the total number of observations is $h = \Sigma_{i,j} h_{ij\cdots z\omega}$. As before $y_{ij\cdots z\omega} \sim N[E(y_{ij\cdots z\omega}), \sigma^2]$. The design matrix is

$$X = \begin{bmatrix} X_1 & X_2 & \cdots & X_m \end{bmatrix}$$

and the parameter vectors are

$$p_{u} = (p_{u1}, p_{u2}, \ldots, p_{un})$$

for $u=1,\ldots,m$. It shall be understood that $p_m$ includes the mean $\mu$, as did $p_2$ in section 2.1. Our model is then

$$E(y) = X_1 p_1 + X_2 p_2 + \cdots + X_m p_m$$

(2.2.1)

Further use of this notation is made in this and later sections. A general form for $C_\alpha$, the matrix for the reduced normal equations (r.n.e.) for $p_\alpha$, will be presented in section 2.3. The s.s. for factor $\alpha$, adjusted for the other $m-1$ factors, is then $s_{s_{p\alpha}} = \hat{p}'_\alpha Q_\alpha$, with a general form for $Q_\alpha$ also given in section 2.3.

By calculating $p'_\alpha Q_\alpha$, the error s.s. under (2.2.1) follows in a manner similar to section 2.1. Without loss of generality, let $\alpha = 1$. The unadjusted estimates are obtained by assuming the conditional model,
\[ E(y) = X_{22}p_{2} + X_{33}p_{3} + X_{44}p_{4} + \ldots + X_{mm}p_{m}. \]

For this model, the normal equations (n.e.) are,

\[
(2.2.2) \begin{bmatrix}
X'X & X'X & \ldots & X'X \\
2 & 2 & \ldots & 2 \\
3 & 3 & \ldots & 3 \\
m & m & \ldots & m
\end{bmatrix}
\begin{bmatrix}
p^* \\
p^* \\
p^*
\end{bmatrix} =
\begin{bmatrix}
X'y \\
2 \\
3 \\
m
\end{bmatrix}
\]

Let \( H \) be the coefficient matrix for the unadjusted estimates \( p^* = (p^*_2, p^*_3, \ldots, p^*_m) \) in (2.2.2) and \( H^- \) be a generalized inverse for \( H \).

Then \( \hat{p}^* = H^-y^* \) where \( y^* = (X' y, X'_3 y, \ldots, X'_m y) \) and the error s.s. for (2.2.1) is \( \text{ss}_e = y'y - \hat{y}^* - \hat{p}'^*Q \).

In calculating a \( C \) matrix for \( p^*_1 \) in section 2.3, a \( g \)-inverse for \( H^- \), is also found, and thus \( \text{ss}_e \) follows. Also, should a solution to the full n.e. be required, \( (\hat{p}_1, \hat{p}_2, \ldots, \hat{p}_m) \), these estimates may be found directly. A solution \( \hat{p}_1 \) is already found from the r.n.e. for \( p^*_1 \).

Then for \( u=2, \ldots, m \), let \( Z_u = X'u'Y - X'u'X_{u1} \hat{p}_{u1} \), and \( Z' = (Z'_2, Z'_3, \ldots, Z'_m) \).

The following estimates are solutions to the full n.e., since they are adjusted for \( \hat{p}_1 \),

\[ (\hat{p}_2, \hat{p}_3, \ldots, \hat{p}_m)' = H^- Z. \]
Since an $H^{-1}$ matrix required for the analysis is calculated in finding a $C_1$ matrix in the manner about to be described, much may be said in favor of the general form for the $C_1$ matrix to be given in the next section.

2.3 A General Form For Calculating $C_\alpha$, The Coefficient Matrix In The Reduced Normal Equations For $p_\alpha$

There are certain advantages in finding a coefficient matrix, $C_\alpha$: (1) the variances and covariances of estimable functions are easier to find, since a generalized inverse is easier to find for the reduced normal equations; (2) the s.s. due to a certain factor can be obtained easily and directly, rather than fitting the full and conditional models; (3) for additive models, the rank of the reduced form as one of several methods (Chakrabarti, 1963) determines the number of estimable functions available; (4) for some interaction models, the main effect and interaction effect estimates may be easier to find.

Without loss of generality, we shall develop a general form for $C_1$, the coefficient matrix for $p_1$. Let us generalize on developments (2.1.6) to (2.1.11). Let $\Delta$ be a linear unbiased estimate for some linear function in $p_1$ only, i.e. $E(\Delta)$ does not contain any other factor effects. We represent $\Delta$ then as

$$\Delta = q_1'X'y + q_2'X'y + q_3'X'y + \ldots + q_m'X'y$$
where as before, $q_i$ are vectors of real numbers.

$$q_i' = (q_{i1}, q_{i2}, \ldots, q_{in_i}) \quad i=1, \ldots, m.$$  

We have already found $E(\Delta)$ for $m = 2$, and calculated $C_i$. It follows then for general $m$,

$$E(\Delta) = (q'X' + q'X' + \ldots + q'n_mX') X p_{11}$$

$$+ (q'X' + q'X' + \ldots + q'n_mX') X p_{22}$$

$$+ (q'X' + q'X' + \ldots + q'n_mX') X p_{m-m}.$$  

If $E(\Delta)$ is to be free of $p_{il}, 2 \leq i \leq m$, the following equations must be satisfied regardless of the true value of $p_{il}, 2 \leq i \leq m$,

$$q'H_{11} + q'H_{22} + \ldots + q'H_{m} = 0$$

where

$$H_{\ell t} = X'X' = (h_{\ell t}) \quad 2 \leq \ell, t \leq m$$

with

$$h_{\ell t} = \Sigma \Sigma \Sigma \Sigma k \ldots k \lambda \ldots n \tau \ldots z$$

and $\lambda, \tau$ being the subscript in the $\ell$ and $t$ position, respectively $(\lambda = 1, 2, \ldots, n_\ell; \tau = 1, 2, \ldots, n_t)$. Note that $H_{\lambda \lambda}$ are diagonal matrices.

For the sake of notation, let $H = (H_{\ell t}), 2 \leq \ell, t \leq m$.

From (2.3.2) we then have
(2.3.3) \[ (q'_2, \ldots, q'_m) = -q'_1 (H_{22} \ldots H_{2m}) \]

or

(2.3.4) \[ (q'_2, \ldots, q'_m) = -q'_1 (H_{12} \ldots H_{1m}) \]

where a generalized inverse for \( H \) is written as

(2.3.5) \[ \left[ \begin{array}{ccc} H & \cdots & H \\ 22 & \cdots & 2m \\ \vdots & \ddots & \vdots \\ H & \cdots & H \\ m2 & \cdots & mm \end{array} \right] = \left[ \begin{array}{ccc} M & \cdots & M \\ 22 & \cdots & 2m \\ \vdots & \ddots & \vdots \\ M & \cdots & M \\ m2 & \cdots & mm \end{array} \right] \]

with \( M \) of order \( n_x \times n_t \). The constant vectors \( q'_{i} \), \( i \geq 1 \), are then

(2.3.6) \[ q'_i = -q'_1 (H_{12}M_{2i} + H_{13}M_{3i} + \ldots + H_{1m}M_{mi}) \]

From equation (2.3.1) in connection with (2.3.6) we obtain

(2.3.7) \[ E(\Delta) = (q'_1 X_{11} + q'_2 X_{12} + \ldots + q'_m X_{1m}) X_{1-1} \]

\[ = q'_1 (H_{11} - H_{12}M_{21} - H_{13}M_{31} - \ldots - H_{1m}M_{m1}) X_{1-1} \]

so that, similar to (2.1.15) \( E(\Delta) \) can be written as
As in section 2.1, the reduced normal equations for $p$ are given by

$$C_{1-1}^{-1} p = q_{1}^{-1} E(\Delta) = q' C_{1} p$$

with

$$C = H_{11} - (H_{12} H_{22} + H_{13} H_{32} + \ldots + H_{1m} H_{m2}) H_{21}$$

$$- (H_{12} H_{23} + H_{13} H_{33} + \ldots + H_{1m} H_{m3}) H_{31}$$

$$- \vdots$$

$$- (H_{12} H_{2m} + H_{13} H_{3m} + \ldots + H_{1m} H_{mm}) H_{m1}$$

or

$$C_{1-1}^{-1} p = q_{1}^{-1} E(\Delta) = q' y + q' y + \ldots + q' y$$

or

$$C_{1-1}^{-1} p = q_{1}^{-1} E(\Delta) = q' \{y - (H_{12} H_{23} + H_{13} H_{33} + \ldots + H_{1m} H_{m3}) y \}$$

As in section 2.1, the reduced normal equations for $p$ are given by

$$C_{1-1}^{-1} p = q_{1}^{-1} E(\Delta) = q' \{y - (H_{12} H_{23} + H_{13} H_{33} + \ldots + H_{1m} H_{m3}) y \}$$

and

$$\Delta = q' \{y - (H_{12} H_{23} + H_{13} H_{33} + \ldots + H_{1m} H_{m3}) y \}$$

and

$$\Delta = q' Q_{1}$$
with
\[(2.3.14) \quad Q_1 = Y_1 - (H_{12} H_{13} \ldots H_{1m}) H_2^{-1} Y_2 \ldots Y_m.\]

In order to calculate \(C\) and \(Q\) one therefore needs to find a generalized inverse for \(H\). Also, to complete the analysis, a generalized inverse for \(C\) is needed. Therefore, methods for computing generalized inverses are given in the following section.

2.4 Calculating Generalized Inverses

Among the many classes of generalized (g-inverse) or pseudo-inverses best summarized by Rhodes (1964) and later by Rao (1967) we shall consider only the "weak" form.

Definition 2.1.: A g-inverse of a matrix \(G\) of order \(k \times k\) is a matrix of order \(k \times k\) denoted by \(G^-\), such that for any \(y\) for which \(Gx = y\) is consistent, \(x = G^- y\) is a solution.

For our purposes if \(G^-\) is a g-inverse for \(G\) then \(GG^- G = G\) will hold and conversely (see Rao, 1967).

The next two theorems are quite programmable and are compatible with section 2.3.

Theorem 2.1 Let \(B\) be a singular square symmetric matrix, \(k \times k\), that may be partitioned into

\[
B = \begin{bmatrix}
B_{11} & B_{12} \\
11 & 12 \\
B_{21} & B_{22} \\
21 & 22
\end{bmatrix}
\]

where \(B_{11}\) is \(k \times k\), \(B_{12}\) is \(k \times k\), \(B_{21}\) is \(k \times k\) and \(k + k = k\).
If \( B_{11}^{-} \) is a g-inverse for \( B_{11} \), then a g-inverse for \( B \) is

\[
B = \begin{bmatrix}
M_{11} & M_{12} \\
M_{12} & M_{22}
\end{bmatrix}
\]

where

\[
M_{22} = (B_{22} - B_{22}B_{11}^{-}B_{12})^{-1}
\]

\[
M_{12} = -B_{12}B_{22}^{-1}M_{22} = M_{22}'
\]

\[
M_{11} = B_{11}^{-} - B_{11}^{-}B_{11}B_{11}M_{11}.
\]

This theorem is taken directly from Rhodes (1964) and has great value in inverting \( H \) in (2.3.10) and (2.3.5), since \( H_{\ell \ell} \), \( \ell = 2, \ldots, m \), are non-singular matrices and hence \( H_{\ell \ell}^{-} = H_{\ell \ell}^{-1} \). For calculating \( M_{11} \) in the above theorem we may repeat the use of this theorem until a unique inverse is attained or use the following theorems.

Theorem 2.2: If \( B \) is a singular \((n \times n)\) matrix of rank \( n - 1 \), then a generalized inverse for \( B \) is

\[
(B + aJ_{nn})^{-1}
\]

where \( a \) is any non-zero real number and \( J_{nn} \) \((n \times n)\) is a unity matrix.

A proof for this theorem is given essentially by Shah (1959) (see also Hinkelmann, 1968a). This latter theorem is most appealing, since practical considerations often require that all contrasts in desired
factor levels be estimable, which means that \( C \) and often other matrices used in calculating \( C \), be of rank one less than their order.

Also, this theorem permits simple computer programming. Another theorem that lends itself to programming is the following given by Bose (1958).

**Theorem 2.3:** Let \( X \) be a design matrix for a multidimensional design of factors \( F_1, F_2, \ldots, F_m \), with factor \( F_i \) having \( n_i \) levels. If the degrees of freedom associated with the \( i \)-th factor is \( n_i - 1 \), then a generalized inverse for \( X'X \) is \((X'X + P)^{-1}\) where

\[
P = \text{DIAG}(\Theta_1, J_{n_1 n_1}, \Theta_2, J_{n_2 n_2}, \ldots, \Theta_m, J_{n_m n_m})
\]

with \( \Theta_1, \Theta_2, \ldots, \Theta_m \) real numbers, \( J_{n_1 n_1} \) are \( n_1 \times n_1 \) unity matrices.

Theorems 2.1, 2.2, and 2.3 are particularly useful for inverting the coefficient matrix for the r.m.e. This matrix will be used in analysis for \( m \) factors, with and without interaction present among them. The latter case (without interaction) has essentially been presented in this chapter. The former case will now be discussed.
CHAPTER III

THE ANALYSIS OF DESIGNS WHERE SOME TWO-FACTOR INTERACTIONS ARE ASSUMED

An analysis of incomplete designs where no interaction is assumed was presented in Chapter II. This analysis can be applied to all such incomplete designs, and therefore applies to those cases where the works of Potthoff (1958), Bose and Srivastava (1964b) and Anderson (1968) do not apply.

Using the approach to solving the normal equations presented in section 2.3, we shall now present an analysis of experimental plans where certain types of two-factor interactions are assumed. An example was given in Chapter I where a price-juice interaction was suspected. The model in this case would contain a vector of parameters for the effects of the price-juice interaction.

As in Chapter I, our attention will be focused on those plans where not all factor level combinations are present. These plans are then incomplete, i.e. are fractional factorials. Further, these fractions can be both asymmetric and irregular, having factors of different levels and having these levels occur with non-proportional frequencies.

Bose and Srivastava (1964a) discuss an approach to the analysis of irregular fractions of $2^{k_1} \times 3^{k_2}$ factorials. However, this work is difficult to apply, and the analysis of irregular fractions with levels greater than three has not been discussed. Potthoff (1958, 1963), Causey (1967), and Anderson (1968) considered the analysis of asymmetric fractions.
However, their work concentrated on designs having partial balance on the factor levels involved. We will investigate the analysis of designs where balance on factor levels is not necessarily present.

Further, the analysis presented in this chapter provides for the possibility of investigating the main effects in the presence of certain two-factor interactions. By presenting this analysis, we do not intend to imply that such analysis shall always be important, but just note here that for some types of interaction, it is indeed useful to investigate main effects. In those cases where it is not useful to do so one investigates "simple" main effects (Cochran and Cox, 1957).

This chapter shall discuss the following: section 3.1 shall present an analysis of a two-factor model with an assumed two-factor interaction and shall introduce the analysis of m factor models with two-factor interactions; section 3.2 presents an analysis of an m-factor model with one interaction; section 3.3 discusses models with two interactions; and section 3.4 discusses the analysis for models with three or more two-factor interactions. Also discussed in section 3.4 is a "weakness" of the approach to the analysis presented in this chapter.

3.1 The Two-Factor Design With Interaction

The analysis of a two-factor model with interaction present for an incomplete design is presented by Scheffé (1959). However, we shall just mention these results and introduce notation that shall serve us later. Our model is now
(3.1.1) \[ E(\bar{y}_{ij}) = p_{ij}, \quad i = 1, \ldots, n; \quad j = 1, \ldots, n, \quad \omega = 0, \ldots, h_{ij}; \]

where \( p_{ij} = \mu + p^i_1 + p^j_2 + t^{ij}_{12} \) and \( t^{ij}_{12} \) represents the effect of interaction between factors one \((F_1)\) and two \((F_2)\) at levels \( i \) and \( j \) respectively.

In matrix notation our model can be written as

(3.1.2) \[ E(y) = X_{(12)} p_{(12)} \]

Where \( y \) is \( h \times 1 \) and \( h = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} h_{ij} \). If \( w \) is the number of level combinations of \( F_1 \) and \( F_2 \) that appear in \( y \), then \( p_{(12)} \) is \( w_{12} \times 1 \).

We shall present \( p \) as if all \( n_1 \times n_2 \) level combinations do appear, with the understanding that those level combinations which do not appear shall be deleted from \( p \). This vector is then

\[ p'_{(12)} = (p^{11}_{12}, p^{12}_{12}, \ldots, p^{n_1 n_2}_{12}, \ldots, p^{n_1 n_2}_{12}) \]

and \( X_{(12)} \) is the design matrix, \( h \times w_{12} \), for \( p \). The parameter vector \( t \) then is defined as \( t'_{(12)} = (t^{11}_{12}, t^{12}_{12}, \ldots, t^{n_1 n_2}_{12}, \ldots, t^{n_1 n_2}_{12}), 1 \times w \).

For the corresponding sum of squares (s.s.) for error and for the parameter \( p \), we will use the notation introduced in Chapter II, \( s s_e \) and \( s s \) respectively. The s.s. due to the parameter vector \( t \) will be expressed as \( s s_{t_{(12)}} \).

Let \( H_{(12)(12)} = X'_{(12)} X_{(12)} \) and \( X'_{(12)} Y = Y_{(12)} \). Then we may express the s.s. due to \( p_{(12)} \) as
Note that this s.s. is the s.s. due to the combined main effects and interaction effects of factors one and two. If the conditional model is (2.1.2) the s.s. due to $t_{12}$ is

$$ s_{t_{12}} = s_{ce} - s_{e} $$

$$ = (y'y - p_1^\prime Q_1 - y_2^\prime H_2^{-1} y_2) - (y'y - s_{p_{12}}) $$

$$ = y_{12}^\prime (H_{12})^{-1} (12) - y_{12}^\prime H_{22}^{-1} y_{12} - \hat{p}_1^\prime Q_1 $$

with $w_{12} = n_1 - n_1 + r_1$ d.f.

The S. S. Due to Main effects and Interaction

If the d.f. for $s_{t_{12}}$ are $(n_1 - 1)(n_2 - 1)$ or if $w_{12} = n_1 n_2$ one may use the approach of Scheffe (1959, p. 118), or proceed as Potthoff (1958, 1963) and Anderson (1968). The s.s. due to main effects is then found from the s.s. due to contrasts in the following parameters, $f_{i.}$ and $f_{.j}$:

$$ f_{i.} = \hat{p}_{12} - \hat{p}_{12}^\prime - \hat{p}_{12} $$

(3.1.3)

$$ f_{.j} = \hat{p}_{12}^\prime - \hat{p}_{12} $$

$$ = y_{12}^\prime (H_{12})^{-1} (12) - y_{12}^\prime H_{22}^{-1} y_{12} - \hat{p}_1^\prime Q_1 $$

with $w_{12} = n_1 - n_1 + r_1$ d.f.
where

\[ \bar{p}_{12}^i = \frac{1}{n_2} \sum_{j=1}^{n_2} p_{12}^{ij} \]

\[ \bar{p}_{12}^j = \frac{1}{n_1} \sum_{i=1}^{n_1} p_{12}^{ij} \]

\[ \bar{p}_{12}^{..} = \frac{1}{n_1n_2} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} p_{12}^{ij} \]

Note that if one applied the restrictions

(3.1.4) \[ \sum_{i=1}^{n_1} p_{12}^i = \sum_{j=1}^{n_2} p_{12}^j = \sum_{i=1}^{n_1} t_{12}^{ij} = \sum_{j=1}^{n_2} t_{12}^{ij} = 0, \]

then \[ f_{1.} = p_{12}^1 \] and \[ f_{.j} = p_{12}^j. \] Potthoff (1958, p.60) proved that the expression for the s.s. due to contrasts in \[ f_{1.} \] and in \[ f_{.j} \] is the same as the expression given by Scheffe (1959, p.118) for the s.s. due to main effects. Also, the s.s. due to contrasts in the following parameters,

(3.1.5) \[ f_{12}^{ij} = p_{12}^{ij} - \bar{p}_{12}^i - \bar{p}_{12}^j + \bar{p}_{12}^{..} \]

when \( w = n_1n_2 \) is the s.s. due to \( t_{12}^{ij} \) since by applying (3.1.4), \[ f_{12}^{ij} = t_{12}^{ij}. \]

Now if the d.f. for \( s_{12}^{i} \) are not \( (n-1)(n-1) \), or if \( w \neq n_1n_2 \) we can proceed to find a test on main effects as suggested by Elston and Bush (1961). Elston and Bush (1961) indicated how some comparisons on the true cell means, i.e. \( E(y_{12}^{ij}) \), can be used such that tests on first order interaction and on main effects in the presence of assumed
interaction can be made. This procedure amounts to testing main effects when there are no observations in some assemblies. We shall refer to Elston and Bush (1961) later in discussing m-factor models.

The Analysis of m-Factor Models

An analysis summary of m-factor models with interaction present will first be given in order to better acquaint the reader with later developments. The following models shall be referred to:

\[(1) \quad E(y) = \sum P_{12} + X_{33}P_3 + X_{44}P_4 + \ldots + X_{m-m}P_{m-m}\]

\[(2) \quad E(y) = \sum P_{12} + X_{34}P_{34} + X_{55}P_5 + \ldots + X_{m-m}P_{m-m}\]

\[(3) \quad E(y) = \sum P_{12} + X_{13}P_{13} + X_{44}P_{44} + \ldots + X_{m-m}P_{m-m}\]

\[(4) \quad E(y) = \sum P_{12} + X_{13}P_{13} + X_{23}P_{23} + X_{44}P_{44} + \ldots + X_{m-m}P_{m-m}\]

In each case we solve the r.n.e. for \(p_{12}\) in order to obtain \(C_{12} \ldots 12\) and an estimate \(\hat{p}_{12}\). The s.s. due to \(p_{12}\) (not \(t\)) is then \(SS = p_{12}^{-12}Q_{12}\).

The unadjusted estimates for \(p^*\) are found as in section 2.2., and the unadjusted s.s., \(Y^*H Y^*\), then follows. (The nature of \(p^*\), \(Y^*\), and \(H\) will change with each model). The s.s. due to error in each case will be expressed as \(y'y - Y^*H Y^* - \hat{p}_{12}^{-12}Q_{12}\).

Some interaction s.s. and main effect s.s. can be found either by subtraction or from the s.s. due to contrasts in "f" parameters similar
to (3.1.3) defined in terms of $p_{12}^{ij}$. Other interaction and main effect s.s. can be found from the s.s. due to contrasts involving parameters $p_{23}^{ik}$, or $p_{23}^{jk}$, etc. In each case the corresponding estimates are found from a set of reduced normal equations. Without loss of generality we shall present estimates for the effects involving $F_1$ and $F_2$. Other estimates may be found similarly. The corresponding s.s. will be presented for each model for the two cases: (a) when all the d.f. for ss are available and (b) when not all the d.f. for ss are not available. The d.f. available for ss may be checked either by inspecting the rank of $C_{12}$ or by methods of Chapters IV and V.

For the analysis of models (3) and (4), we will present a reduced form of the normal equations for $p_{12}$ for each level, i, of $F_1$. The intent here is to simplify the solution to the normal equations.

3.2 The $m$-Factor Design With One Interaction

The Analysis of Model (1)

Assume that we have an $m$-factor incomplete design where, without loss of generality, interaction between factors one and two may be present. A typical observation is represented by $y_{ijk...z}$ where the notation of section 2.2 applies to our observations and the parameters $p_{u}$, $l = 1,...,n_u$, $u = 1,...,m$. The model is now

\[(3.2.1) \quad E(y_{ijk...z}) = p_{12}^{ij} + p_{3}^{k} + \ldots + p_{m}^{z}\]

where $p_{12}^{ij} = p_{1}^{i} + p_{2}^{j} + t_{12}^{ij}$. Note that $p_{12}^{ij}$ does not include $\mu$ as it did
in section 3.1. Using the notation of 2.2 and 3.1, (3.2.1) may be expressed as model (1):

\[(3.2.2) \quad E(y) = X_{(12)}p_{12} + X_{3}p_{3} + \ldots + X_{m}p_{m} \]

The above model is similar to an \((m-1)\)-factor model with no interaction. Therefore, the analysis follows that of Chapter II. Let \(X'_{(12)}X_{(12)} = H_{(12)(12)} \) and \(X'_{(12)}X_{u} = H'_{(12)u} = H'_{u(12)}, \) for \(u = 3, \ldots, m.\)

The reduced normal equations (r.n.e.) for the parameter \(p_{12}\) are then

\[
C_{12}^{-1}p_{12} = Q_{12}^{-1}
\]

where \(C_{12}\) is similar to \((2.3.10)\) and \(Q_{12}\) is similar to \((2.3.14)\).

Note that the matrix \(H\) in \((2.3.10)\) is now

\[(3.2.3) \quad H = \begin{bmatrix}
H_{33} & H_{34} & \ldots & H_{3m} \\
H_{m3} & H_{m4} & \ldots & H_{mm}
\end{bmatrix}
\]

\(C_{(12)}\) is \(w_{12} \times \frac{w}{12}\) of rank \(w_{12} - r_{12}\), where \(r_{12}\) is an integer, \(1 \leq r_{12} \leq w_{12} - 1\). The s.s. due to \(p_{12}\) is \(ss_{12} = p'_{12} Q_{12}^{-1} p_{12}\), with d.f. = \(w_{12} - r_{12}\).

The S.S. due to Main effects And Interaction

If the d.f. available for ss are \(n_n - 1\), i.e. \(w_{12} - r_{12} = n_n - 1\)
and, or equivalently, if the d.f. for $ss_{t}$ is $(n - 1)(n - 1)$, we can define the parameters $f_{ij}$, $f_{i.}$, and $f_{.j}$ as in (3.1.3) and (3.1.5).

The s.s. for interactions, $ss_{t}$, follows from the s.s. due to contrasts in $f_{ij}$. The s.s. for main effects under the presence of interaction follows from contrasts in $f_{i.}$ and $f_{.j}$. For example, an estimate of

\[
\begin{align*}
\hat{t}_{11} + \hat{t}_{22} - \hat{t}_{21} - \hat{t}_{12} &= L't \\
\text{is then } \hat{f}_{11} + \hat{f}_{22} - \hat{f}_{21} - \hat{f}_{12} \text{ or alternatively,} \\
p_{11} + p_{22} - p_{21} - p_{12} &= L'\hat{p} \\
\text{The s.s. due to } L't \text{ is then} \\
(3.2.4) \\
(L'\hat{p})' \cdot [L' C^{-1} L]^{-1} (L'\hat{p})
\end{align*}
\]

where $C^{-1}$ is a g-inverse for $C$. The main effect contrasts in $p_{1.}$ and $p_{.2}$, and their corresponding s.s. are also found from contrasts in $p_{12}$.

Appendix I presents a numerical example of such an analysis.

If the d.f. available for $ss_{t}$ is not $n n - 1$ but are $w - r \neq 1 2 12 12$ $n n - 1$, or equivalently, the d.f. for $ss_{t}$ are not $(n - 1)(n - 1)$, there still may be comparisons on two-factor interactions and main effects that can be tested. First, it still may be possible to define meaningful effects in terms of the "f" parameters, $f_{ij}$, $f_{i.}$, and $f_{.j}$, as in section 3.1. If so, then the s.s. due to individual contrasts may be found as before. However, if it is not possible to define such meaningful contrasts, it still may be able to compare factor level combinations as do Elston and Bush (1961).

Further, if $w - r \neq n n - 1$, the s.s. due to $t_{-12}$ can be obtained by subtraction as in section 3.1. Let the error s.s. under the
conditional model given by (2.2.1) be $\text{ss}_e^{(1.2.3)}$. Let the error s.s. from (3.2.2) be denoted by $\text{ss}_e^{(12.3)}$. From a solution to the r.n.e. for $p_{12}^\ast$, $\text{ss}_e^{(12.3)} = y'y - Y^*H^* - p_{12}^\ast Q$ with d.f. $= h - w + r - r$, where $H$ is of rank $r$, and is given in (3.2.3), and $Y^*H^* - Y^*$ is the unadjusted s.s. found as in section 2.2. Then $\text{ss}_t_{12} = \text{ss}_e^{(1.2.3)} - \text{ss}_e^{(12.3)}$.

Testing the absence of $F$ and Interaction

Also using the conditional model $E(y) = Xp_{2} + Xp_{3} + \ldots + Xp_{m} = Xp_{i}$ and (3.2.2) one could test $p_{1} = 0$ and $t_{12} = 0$, or simply the absence of factor one and of the first order interaction. This may be especially appealing if one would desire to test the main effects of factor one when not all level combinations of factors one and two are present.

Let the error s.s. under conditional model be denoted by $\text{ss}_e^{(2.3)}$ and the error s.s. under (3.2.2) be denoted by $\text{ss}_e^{(12.3)}$. Then the s.s. due to $p_{1} = 0$ and $t_{12} = 0$ would be $\text{ss}_{p_{1}, t_{12}} = \text{ss}_e^{(2.3)} - \text{ss}_e^{(12.3)}$.

The d.f. for $\text{ss}_{p_{1}, t_{12}}$ are also received by subtraction. A similar procedure can be used for testing the absence of $F$ and interaction.

This approach may be needed later in discussing main effects which are defined by more parameters than $p_{12}^{1j}$, since using our form of the reduced normal equations for $p_{12}^*$ only, and corresponding g-inverse $G_{12}^*$, we do not have the variance-covariance matrix for other parameters. More shall be said on this in section 3.4.
The Analysis of Factors Other Than $F_1$ and $F_2$

The analysis on factors one and two is now complete. For individual tests on factors three to $m$, one can obtain a set of reduced normal equations for each factor, solve these equations and obtain the respective s.s. for each factor. For example, to adjust $p_3$ for the parameters $p_{-12}, p_{-4}, \ldots, p_m$, a coefficient matrix for the reduced normal equations is found by interchanging the integer subscripts (12) and 3 in the expression for $C_{12}$.

Note that since $\nu$ is associated with $p_m$, we redefine our model so that $\nu$ is associated with $p_{m-1}$ before the matrix $C_m$ is calculated for factor $m$. This terminates our discussion of $m$ factor models with one interaction assumed.

3.3 The $m$-Factor Design With Two Interactions

The Analysis of Models (2) and (3)

For models of the form $E(y_{ijkl\ldots za}) = p_{12}^{ij} + p_{34}^{k\ell} + \ldots + p_m^z$, model (2), where two first order interactions are present and no factor interacts with another more than once, the analysis follows the methods of Chapter II, and sections 3.1 and 3.2. The analysis of model (3) where a factor interacts with two other factors given by

\[(3.3.1) \quad E(y_{ijkl\ldots za}) = p_{12}^{ij} + p_{13}^{jk} + p_{4}^{\ell} + \ldots + p_m^z\]

shall be discussed in this section. Without loss of generality we discuss this in terms of model (3.3.1). The parameters are similar to
those defined in section 3.2 with the exception that now

\[ p_{ij} = \frac{1}{4} p_{12}^i + p_{12}^j + \frac{1}{k} p_{13}^i + t_{12} \]

\[ p_{ik} = \frac{1}{4} p_{12}^i + p_{12}^k + t_{12} \]

Let \( w \) be the number of factor level combinations from \( F \) and \( F \) that actually occur. Let \( p_{13} = (p_{11}^1, p_{12}^1, \ldots, p_{1n_3}^1, \ldots, p_{1n_3}^{n_3}) \) be of order \( 1 \times w \), and \( X_{(13)} \) the \( h \times w \) corresponding design matrix for \( p_{13} \). Then, with \( p_{-12} \) and \( X_{(12)} \) as defined before (3.3.1) may be expressed as

\[ E(y) = X_{(12)}p_{-12} + X_{(13)}p_{-13} + X_{4}p_{4} + \ldots + X_{m}p_{m} \]

This model is somewhat similar to an \( (m-2) \)-factor model with no interaction present and the analysis of this model, then, will follow procedures similar to previous development. However, there will be some different steps taken in our approach.

For a fixed level of \( F, i \), we will find an expression for the coefficient matrix of the estimate \( \hat{p}_{12} \) in the r.n.e. for \( F \). This coefficient matrix shall be expressed as \( C_{12}^{i} \). The reduced normal equations for \( p_{-12} \) can then be expressed as sets of equations \( C_{12}^{i}p_{-12} = Q_{12}^{i}, i = 1, \ldots, n \), where the vectors \( p_{-12} \) and \( Q_{12}^{i} \) are similar to those defined in section 2.3. (We shall define these vectors explicitly in the following developments.) An expression for this coefficient matrix, \( C_{12}^{i} \), shall be given first. We need to define the matrices involved. Let \( X_{(12)} \) be a
design matrix that is obtained from $X_{(12)}$ by fixing the $i$-th level of factor $F_i$. Note that $X_{(12)}$ is $h \times w_{12}$ and

$$X_{(12)} = (X_{(12)}^{(1.)}, X_{(12)}^{(2.)}, \ldots, X_{(12)}^{(n_{12})})$$

and $X_{(12)}^{(i.)}$ is $h \times n_{12}^i$ where $n_{12}^i$ is the number of levels of factor $F_i$ that appear with the $i$-th level of factor $F_i$, and $\Sigma n_{12}^i = w_{12}$. Then

$$H_{(i.)}(12) = X_{(12)}(i.)' X_{(12)}(i.)$$

is an $n_{12}^i \times n_{12}^i$ diagonal matrix with elements

$$h_{i1}, h_{i2}, \ldots, h_{in_{12}}^i$$

where

$$h_{ij} = \Sigma \Sigma h_{ik} \Sigma \Sigma z_{ijkl\ldots z},$$

and $X_{(12)}(i.)' X_{(12)}(i.) = H_{(i.)}(12)(12)$, $X_{(12)}(i.)' X_{(12)}(i.) = H_{(i.)}(12)(12)$ for $u = 4, \ldots, m$. Let $X_{(12)}' X_{(12)} = H_{(12)}$, $u, v = 4, \ldots, m$, $X_{(13)}' X_{(13)} = H_{(13)}$, $u, v = 4, \ldots, m$, then

$$X_{(12)}' X_{(12)} = H_{(12)} = \begin{bmatrix}
H_{(13)(13)} & H_{(13)4} & \ldots & H_{(13)m} \\
& & & \\
& & \ddots & \\
& & & \ddots \\
H_{m(13)} & H_{m4} & \ldots & H_{mm}
\end{bmatrix}$$

(3.3.3) $H = \begin{bmatrix}
H_{(13)(13)} & H_{(13)4} & \ldots & H_{(13)m} \\
& & & \\
& & \ddots & \\
& & & \ddots \\
H_{m(13)} & H_{m4} & \ldots & H_{mm}
\end{bmatrix}$

Given a g-inverse $H^{-1}$, the r.n.e. for $F$ with $i$ fixed, follow in a manner similar to section 2.3 and the coefficient matrix is calculated in a form similar to (2.3.10).

A formal presentation of the development for $C_{12}^i$, is because of the notation involved, difficult to follow. However, this development is
quite similar to that in section 2.3 and is therefore, not really necessary. From (2.3.10), fixing the $i$-th level of $F_1$, we have the expression for $C_{12}^i$,

$$C_{12}^i = H(1)_{(12)}H(1)_{(13)}...H(1)_{(12)m}H^{-1}H(1)_{(13)}H(1)_{(12)m}...H(1)_{(12)m}.$$

$C_{12}^i$ is a $n_{12}^i \times n_{12}^i$ matrix of rank $n_{12}^i - r_{12}^i$, where $r_{12}^i$ is an integer function of $i$. $C_{12}^i$ is a coefficient matrix for $p_{12}^i$, where $p_{12}^i$ is given by

$$p_{12}^i = (p_{12}^{11}, p_{12}^{12}, p_{12}^{13}, ..., p_{12}^{in_{12}^i}).$$

The right side of the r.n.e. is similar to (2.3.14)

$$Q_{12}^i = X(1)_{(12)}X^{-1}H(1)_{(12)}H(1)_{(13)}...H(1)_{(12)m}H^{-1}H(1)_{(13)}X_4...X_m')y$$

where corresponding to $p_{12}^i$, $Q_{12}^i$ is defined by

$$Q_{12}^i = (Q_{12}^{11}, Q_{12}^{12}, Q_{12}^{13}, ..., Q_{12}^{in_{12}^i}).$$

A solution to the r.n.e. is found for each $i$, $i=1, ..., n$, and a parameter $p_{12} = (p_{12}^1, p_{12}^2, ..., p_{12}^{n_{12}^1})$ is now available for calculating the corresponding s.s. for testing $p_{12} = 0$,

$$ss_{p_{12}} = \sum_{i=1}^{n_{12}^i} p_{12}^i Q_{12}^i, \text{with } \chi(n_{12}^i - r_{12}^i) \text{ d.f.}$$

Note that a g-inverse for $C_{12}^1$ is found easily, since $C_{12}^1 = \text{diag}(C_{12}^1, C_{12}^2, ..., C_{12}^n)$.

Let $ss_{e(12,13)}$ be the error s.s. under the
model (3.3.2) and \( \text{ss}_e(13.4) \) the error s.s. under the model

\[
E(y) = X_{(13)}p_{13} + X_4p_4 + \ldots + X_mp_m.
\]

This model is similar to (3.2.2) and may be considered a "conditional" model. Then \( \text{ss}_e(12.13) = \text{ss}_e(13.4) - \text{ss}_p \). Note \( \text{ss}_e(13.4) \) is easy to calculate, since it uses the \( H \) matrix of (3.3.3); \( \text{ss}_e(13.4) \) is calculated in a manner similar to approach given in section 2.2. Let the design matrix \( X = [X_{(13)}, X_4, \ldots, X_m] \) of (3.3.4) have rank \( a \), then

\[
h = a - \sum_{i=1}^{n_1} \left( n_{i1} - r_{i1} \right) \text{ are the d.f. for } \text{ss}_e(12.13).
\]

The S.S. due to Main Effects of \( F_2 \) and Interaction

If the d.f. available for \( \text{ss}_p \) are \( n(n-1) \), i.e., if the rank of \( C \) is \( \sum_{i=1}^{n_1} \left( n_{i1} - r_{i1} \right) = n(n-1) \), the main effects of factor two and the interaction effects of factors one and two may be received from

\[
f_{.j} = \bar{p}_{.j} - \bar{p}^3
\]

\[
f_{i.j} = p_{i.j} - \bar{p}_{.j} - \bar{p}_{i.} + \bar{p}^3
\]

where \( \bar{p}_{.j} \), \( \bar{p}_{i.} \), and \( \bar{p}^3 \) are defined as in section 3.1. Note that by applying (3.1.4), \( f_{.j} = p^j_2 \) and \( f_{i.j} = t^i_j \). Then the s.s. contrasts of factor two of or the two-factor interaction may be received by (3.2.4).
If the d.f. available for ss are not \( n(n - 1) \), then we proceed, as we did for the similar case in section 3.2. Some meaningful effects may still be defined in terms of \( f_{.j} \) and \( f_{1j} \), and the s.s. for these may be found by (3.3.4). Another approach could be that of Elston and Bush (1961). By comparing factor level combinations, we still may be able to define some meaningful tests. Further, the s.s. due to (interaction), \( ss_{t} \), can be found as the difference \( ss_{e(13.2)} - ss_{e(12.13)} \), where \( ss_{e(13.2)} \) is the error s.s. under

\[
E(y) = X_{(13)}(13) + X_{(12-1)} + X_{(14-1)} + \ldots + X_{(m-m)}.
\]

The d.f. for \( ss_{t} \) is then \( \sum_{i=1}^{n} [n_{i.} - r_{i.}] - b \), where \( b \) is the rank of the coefficient matrix in the r.n.e. for \( p_{2} \) under (3.3.3). Note that if

\[
\sum_{i=1}^{n} [n_{i.} - r_{i.}] = n(n - 1) \quad \text{and} \quad b = n - 1,
\]

then the d.f. for \( ss_{t} \) are

\[
(n - 1)(n - 1).
\]

The S.S. Due to The Main Effects of \( F_{1} \)

In order to estimate main effect contrasts of \( F_{1} \), we could proceed as follows. From the r.n.e. for \( p_{2} \), an estimate, \( \hat{p}_{-12} \), can be found. Then, using the \( H \) matrix defined in (3.3.3) and the procedures of section 2.2, an estimate \( \hat{p}_{-13} \) can be found. Let the d.f. available for \( F_{1} \) be \( n - 1 \). (This may be checked by methods of Chapter IV). Then the estimates of \( F_{1} \) effects may be define in terms of
If we apply (3.1.4), the $f_i$ are then $p_{1}^{i}$. (The parameters $p_{12}^{i}$, $p_{13}^{i}$ are defined similar to $p_{12}^{j}$ and $p_{13}^{j}$ in section 3.1.) A test in $f_i = 0$ would require the s.s. necessary to test the absence of factor one. Let this s.s. defined by $ss_{f_i}$. Then $ss_{f_i} = ss_{e(2.3)} - ss_{e(12.13)}$

where $ss_{e(2.3)}$ is the s.s. due to error under the conditional model

$$E(y) = X_{2-2}p_{2} + X_{3-3}p_{3} + \ldots + X_{m-m}p_{m}$$

The d.f. for $ss_{f_i}$ would then be $(n - 1) + (n - 1)(n - 1) + (n - 1) = (n - 1)(n + n - 1)$, [if the d.f. for $ss_{p_{12}}$ are $n (n - 1)$, and the d.f. for $ss_{p_{13}}$ (due to $p_{13}$) are $n (n - 1)$].

However, a test on the main effects of $F$ under the presence of interaction would involve a little more work. Note that the variances of these effects cannot be defined in terms of one g-inverse, $C_{12}^{-}$, as in section 3.2. This situation then, illuminates a weakness in the general approach developed in Chapter III for models of the type (3.3.2). (Note that this weakness is only for certain types of main effects.) In order to find the variances of estimates of effects in factor one, we can not proceed as before. However, the s.s. due to these effects still may be expressed. Let $X = [X_{12}, X_{13}, X_{w}, \ldots, X_{m}]$ be the design matrix for model (3.3.2). Let $p' = (p_{12}, p_{13}, p_{4}, \ldots, p_{m})$ and let $L'p$ be a main effect contrast (i.e. defined in terms of $p_{12}$ and $p_{13}$). Then the s.s. due
to this contrast is

\[(3.3.7) \quad (L'\hat{\beta})'\left(L'(X'X)^{-1}L\right)^{-1}(L'\hat{\beta})\]

where \(\hat{\beta}\) is a solution to the normal equations and \((X'X)^{-1}\) is a g-inverse for \((X'X)\). Further, the variance of \(L'\hat{\beta}\) is \(L'(X'X)^{-1}\Sigma^{-2}\). [These last results can be found in Rao, 1967].

If the d.f. available for factor one is not \(n_1 - 1\), we still may be able to make some tests on factor one under the presence of first order interaction. For these tests we may be able to compare factor level combinations as do Elston and Bush (1961), as mentioned earlier.

This essentially covers the analyses of factors one and two under the assumed presence of two first order interactions. A similar approach may be used in the analysis of factors one and three. For tests on \(p_u\), \(u = 4, \ldots, m\), the corresponding r.n.e. for each \(p_u\) may be found and estimates \(\hat{p}_u\) calculated. The necessary s.s. and variances of estimates may then be found from the g-inverses of \(C_u\), \(u = 4, \ldots, m\).

3.4 The M-Factor Design With Three Interactions

The Analysis of Model (4)

The analysis of a design model with three first order interaction parameter vectors, as for example, Model (4):

\[(3.4.1) \quad E(y) = X_{(12)}p_{12} + X_{(13)}p_{13} + X_{(23)}p_{23} + X_{u}p_u + \ldots + X_{m}p_m\]

is similar to what has been presented. The design matrices \(X_{(12)}\), \(X_{(13)}\) and \(X_{u}\), \(u = 4, \ldots, m\) are defined as before, the parameters \(p_u\),
are as before, and the parameters, \( p_{12}, p_{13}, \) and \( p_{23}, w \times 1 \) with design matrix \( X_{(23)}, h \times w_{23}, \) are now defined from the following:

\[
\begin{align*}
  p_{ij}^{12} &= \frac{1}{2} p_i^1 + \frac{1}{2} p_j^1 + t_{ij}^{12} \\
  p_{ik}^{13} &= \frac{1}{2} p_i^1 + \frac{1}{2} p_k^1 + t_{ik}^{13} \\
  p_{jk}^{23} &= \frac{1}{2} p_j^1 + \frac{1}{2} p_k^1 + t_{jk}^{23}
\end{align*}
\]

where \( t_{ik}^{12} \) is the interaction effect of \( F_1 \) and \( F_2 \) at levels \( j \) and \( k, \) and \( w \) different level combinations of \( F_1 \) and \( F_2 \) occur. One then solves the r.n.e. for \( p \) at fixed level, \( i, \) of \( F. \) The \( C_{ij}, \) \( i, \) and \( Q_{ij}, \) matrices and vectors are similar to those presented for model (3.3.2). To find these explicitly, one follows the analysis of sections, 2.3 and 3.3.

The s.s. for testing \( p_{ij}^{12} = 0 \) is then \( \text{ss}_{12} = \sum_i \hat{p}_{i}^{1} Q_{i}. \) This s.s. now corresponds to the s.s. for testing interaction effects (for testing \( t_{ij}^{12} = 0 \)). If the d.f. for \( \text{ss}_{t} \) are \( \sum_{i=1}^{12} (n_i^1 - r_i^1) \), \( (n_i^1 \) and \( r_i^1 \) are defined in section 3.3) and are equal to \( (n - 1)(n - 1) \), then \( \text{ss}_{t} \) is the s.s. due to contrasts on \( f_{ij} = p_{ij}^{12} - p_i^1 - \bar{p}_{j}^{12} + \bar{p}_{.}. \) since applying conditions (3.1.4), \( f_{ij} = t_{ij}^{12}. \) Then one uses procedures given in section 3.2 to find the variances and necessary s.s. due to the interaction contrasts. A similar approach would be used for interaction contrasts involving \( p_{ik}^{13} \) or for contrasts involving \( p_{jk}^{23}. \)

If \( \sum_{i=1}^{12} (n_i^1 - r_i^1) \) are not \( (n - 1)(n - 1) \) (and this may be checked...
using methods of Chapter IV) then $ss_{t}$ may be found by subtraction. Let $ss_{e(13.23)}$ be the error s.s. under

$$E(y) = X_{(13)}p_{13} + X_{(23)}p_{23} + X_{4}p_{4} + \ldots + X_{m-m}$$

and let the error s.s. for (3.4.1) be $ss_{e(12.13.23)}$. Then $ss_{t} =$

$$ss_{e(13.23)} - ss_{e(12.13.23)} \text{ with } \sum_{i=1}^{12} (n - r_{i}) \text{ d.f.}$$

A similar approach follows for interaction between $F$ and $F$ or between $F$ and $F$.

The S.S. Due to Main Effects

For estimates of main effects, we have a situation similar to section 3.3. If the d.f. for $ss_{t}$, $ss_{t}$, and $ss_{t}$, are respectively

$$(n - 1)(n - 1), (n - 1)(n - 1), \text{ and } (n - 1)(n - 1),$$

we can define the main effects of factors one, two, and three in terms of the contrasts

$$f_{1} = \bar{p}_{12} + \bar{p}_{13} - \bar{p}_{1} - \bar{p}_{2}$$

$$f_{2} = \bar{p}_{23} + \bar{p}_{3} - \bar{p}_{3}$$

$$f_{3} = \bar{p}_{12} + \bar{p}_{23} - \bar{p}_{1} - \bar{p}_{3}$$

where $\bar{p}_{12}, \bar{p}_{13}, \bar{p}_{23}, \bar{p}_{3}$, and $\bar{p}_{13}$ follow from before and

$$\bar{p}_{23} = \frac{1}{n} \sum_{k=1}^{n} p_{jk}$$

$$\bar{p}_{23} = \frac{1}{n} \sum_{j=1}^{n} p_{jk}$$
\[ \bar{p}_{23} = \frac{1}{n} \sum_{i=2}^{n} \sum_{j=3}^{k} p_{23}^{jk} \]

Note that applying the restrictions,

\[ \sum_{i=1}^{1} \sum_{j=2}^{12} t_{ij}^{12} = \sum_{i=1}^{1} \sum_{j=12}^{13} t_{ik}^{12} = \sum_{i=1}^{1} \sum_{k=13}^{23} t_{ik}^{12} = \sum_{j=23}^{k} p_{23} = 0, \]

one then defines \( f_{1..} = p_{1} \), \( f_{.j} = p_{2} \), \( f_{..k} = p_{3} \).

We may approach the analysis of these effects as we did in section 3.3. An estimate \( \hat{p}_{12} \) can be found from the r.n.e. for \( p_{12} \). Estimates \( \hat{p}_{13} \) and \( \hat{p}_{23} \) then may be found by the procedure indicated in section 2.2, where using \( \hat{p}_{12} \) and the full normal equations, we find adjusted estimates \( \hat{p}_{12}^{-1} \) and \( \hat{p}_{13}^{-1} \). Therefore estimates \( \hat{f}_{1..} \), \( \hat{f}_{.j} \), and \( \hat{f}_{..k} \) can be found.

As in section 3.2, we may test \( f_{1..} = 0 \), a test for the absence of factor one. But this test would involve \( p_{1} = 0, t_{12} = 0, \) and \( t_{13} = 0, \) which may seem a bit unrealistic. A similar case could be argued for \( f_{.j} = 0 \) or \( f_{..k} = 0. \)

A more realistic approach to testing the main effects under the assumed presence of three first order interactions would obviously require a little more work. As in the case of the main effects of \( F \) in section 3.3., the variances of the main effects for \( F_{1}, F_{2}, \) and \( F_{3} \) cannot be defined in terms of a g-inverse for the coefficient matrix in the r.n.e. (i.e. \( C_{12}^{-}, C_{13}^{-}, \) or \( C_{23}^{-} \)). Therefore, we must proceed as we did in (3.3.7). Let \( X = [X_{(12)} X_{(13)} X_{(23)} X_{4} \ldots X_{m}] \) be the design matrix for model (3.3.2). Let \( p' = (p_{12}, p_{13}, p_{23}, p_{4}, \ldots, p_{m}) \) and let \( L'p \) be a main
effect contrast for factor one [i.e. defined in terms of $p_{-12}$ and $p_{-13}$].

Then the s.s. due to this contrast is

$$(L'\hat{p})' (L'(X'X)^{-L})^{-1} (L'\hat{p})$$

where $\hat{p}$ is a solution to the normal equations and $(X'X)^{-}$ is a g-inverse for $(X'X)$. Further, the variance of $L'\hat{p}$ is $L'(X'X)^{-L} \sigma^2$. A similar approach holds then for main effects of $F_2$ and $F_3$.

If the d.f. available for $s_{st}$ is not $(n_1 - 1)(n_2 - 1)$ or the d.f. for factor one is not $(n_1 - 1)$, we still may be able to make some tests on factor one under the presence of first order interaction. For these tests we may compare factor level combinations as do Elston and Bush (1961). A similar approach would hold for $F_2$ and $F_3$.

This essentially covers then the analysis of the effects of $F_1$, $F_2$, and $F_3$ under the assumed presence of three first order (f.o.) interactions. For designs where more than three f.o. interactions are present, the analysis follows sections 3.1, or 3.2, or 3.3. For example, the analysis of the following model

$$E(y) = X_{(12)}p_{-12} + X_{13}p_{-13} + X_{(23)}p_{-23} + X_{(14)}p_{-14} + X_5p_5$$

would follow this section, with a few adjustments. Here the parameters $p_{-12}$, $p_{-13}$, $p_{-14}$ would now be defined

$$p_{-12} = \frac{1}{3} p_{-1} + p_{-2} + t_{-12}$$

$$p_{-13} = \frac{1}{3} p_{-1} + p_{-3} + t_{-13}$$
\[ p_{-14} = \frac{1}{3} p_{-1} + p_{-4} + t_{-14} \]

Therefore, we shall terminate our approach to the analysis of m-factor models here.

The "Weakness" of Our Approach to the Analysis

A discussion of this approach seems relevant here. In finding the r.n.e. for one parameter, say \( p \), the intent is to simplify the solution to the full normal equations, \( X'Xp = X'y \), where in general, the design model is \( E(y) = Xp \), \( X \) is the design matrix, and vector \( p \) is the parameter of effects. However, for models where \( F \) interacts with two or more factors, the s.s. due to this factor one under the assumed presence of the f.o. interactions cannot be found using \( C^- \) alone. Using our approach, the main effects of \( F \) can be estimated, but again the variances of these effects cannot be found using \( C^- \) alone.

However, there are cases where there is little value in investigating such main effect estimates when two-factor interactions are assumed to exist. As mentioned earlier, these are cases where the analysis of the interactions are more important. Further, for those cases where main effects are to be investigated under the assumed presence of two-factor interactions, the special approach developed in this chapter still has some use.

First, it should be said that these problems of main effect analysis only arise for cases where one factor interacts with two or more factors.
Factors that interact only once cause no such problems. Next, it should be said that the special analysis mentioned in Chapter III provides not only a method of solving the r.n.e. but also provides those matrices necessary for calculating the variance-covariance matrix from the full normal equations, a g-inverse for \((X'X)^{-}\).

For example, consider model (2) or (3.3.2). In order to find the variances of and s.s. due to contrasts in

\[ f_{i..} = \bar{p}_{12} + \bar{p}_{13} - \bar{p}_{12} - \bar{p}_{13} \]

a g-inverse for \(X'X\) is required, where \(X'X\) is defined

\[
X'X = \begin{bmatrix}
H_{(12)(12)} & H_{(12)(13)} & \cdots & H_{(12)m} \\
H_{(13)(12)} & H_{(13)(13)} & \cdots & H_{(13)m} \\
& & \ddots & \vdots \\
& & & H_{mm}
\end{bmatrix}
\]

The matrices in \(X'X\) are defined: \(H_{(12)(12)} = X'_{(12)}X_{(12)}, \ H_{(12)(13)} = X'_{(12)}X_{(13)}, \ H_{(12)u'} = X'_{(12)}X_{u'}, \ u = 4, \ldots, m\). Also, \(X'X\) is symmetric and all the other matrices have been defined in Chapter II and earlier sections of this chapter.

Now refer to Theorem 2.1. Let \(H\) be defined as in (3.3.3). Note that this is a submatrix of \(X'X\). Using the notation of Theorem 2.1, let \(X'X = B, H = B_{22}\), and let the g-inverse \(H^{-}\), calculated for finding \(C_{i.}\) in section 3.3, be \(H^{-} = B_{22}^{-}\). Since \(H^{-}\) may be a non-singular matrix,
the calculation of $M_{11} = (B_{11} - B_{12} B_{22} B_{21})^{-1}$ may be simplified, i.e. $(M_{11})^{-1}$ may be a non-singular matrix. Finally, all the submatrices of $X'X^{-1} = B$ may be calculated, and thus a g-inverse for $(X'X)$ may be found from the calculations necessary for the r.n.e. Therefore, if a g-inverse, $(X'X)^{-1}$ is really necessary, as for example, in the analysis of certain main effects, this g-inverse may still be found from the analysis developed in this chapter.

Our analysis given in this chapter will provide some insight to the construction of designs to be discussed in Chapter V. The theory for such construction will be provided in Chapter IV.
CHAPTER IV
CONNECTING EXPERIMENTAL DESIGNS

Now that we have established an analysis of some types of incomplete multidimensional designs in Chapters II and III, let us proceed to develop techniques which will aid in the construction of the minimal designs mentioned in Chapter I.

In this chapter we shall establish a concept of connected designs when first order (two-factor) interaction is assumed. We shall use this concept to augment designs with the minimum number of factor level combinations (assemblies) necessary to do so.

The concept of a connected design is known and used most frequently in the analysis of incomplete block designs. A connected incomplete block design has treatments associated with blocks in such a manner that all treatment differences are estimable.

4.1 A Literature Review And Summary of Results

Anderson (1968) and Srivastava and Anderson (1970) defined the concept of connectedness in main effect MD's, where levels of one factor are connected in the presence of the other m-1 factors. Their results will be given in section 4.2. From these results we shall proceed to establish a similar concept for level combinations of two-factors, and thus define a connected design with assumed two factor interactions. Using this definition, we shall establish a procedure for finding assemblies to augment designs which are not completely connected. Further, we shall confine our interest to augmenting designs with only the mini-
minimum number of assemblies necessary.

Federer (1956, 1961) presents a procedure for adding new treatments (and thus assemblies) to a design in order to estimate the effects of the new treatments. Federer does use the minimum of necessary assemblies, and is concerned with connecting the levels of one-factor treatments. However, the treatments that he adds to a design are "new" treatments, that did not appear in the original design. We shall be concerned with only connecting factor levels that do appear in the original design.

4.2 Augmenting Designs With Assumed First Order Interactions

Here we shall present the basic definitions and theorems for connected designs. These theorems will be used in developing a procedure for augmenting designs.

**Definition 4.1**

A multidimensional design (MD), T, consists of $h = h(T)$ assemblies (factor level combinations) and an associated linear model $LM(T)$.

Let T be a MD with m factors $F_1, F_2, \ldots, F_m$, where $F_i$ has $n_i$ levels. Let $LM(T)$ be of the form $E(y_T) = X_T \cdot p_T$ where $y_T$ is the observation vector (hx1), $X_T$ is the design matrix (hxM(T)), and $p_T$ is the parameter vector $(M(T) x 1)$ with $p_T' = (p_{i_1}^1, p_{i_2}^1, \ldots, p_{i_q}^q, p_{j_1 j_2}^1, \ldots, p_{j_r j_s}^r)$ and

$$p_a^1 = (p_{a_1}, p_{a_2}, \ldots, p_{a_n})$$

$$p_{bc}^{11} = (p_{bc_1}, p_{bc_2}, \ldots, p_{bc_n})$$

Here $p_a (a = i_1, i_2, \ldots, i_q)$ refers to the factor $F_a$ that does not inter-
act with any other factor, and \( p_{bc} ((b,c) = (j_1, j_2), \ldots, (j_r j_S)) \)

refers to pairs of factors that interact with each other. Accordingly, let \( S_a \) denote the set of all levels of \( F_a \), and let \( S_{bc} \) denote the set of pairs of level combinations of \( F_b \) and \( F_c \). \( M(T) \) is the number of components in \( \Phi_T \). The intention here is to estimate with MD \( T \), parameter contrasts of the form

\[
\ell_a \ell'_a
\]

(4.2.1) \( p_a - p_a \)

and

\[
\ell_b \ell'_b + \ell'_b \ell_c - \ell'_b \ell_c - \ell_b \ell_c
\]

(4.2.2) \( p_{bc} + p_{bc} - p_{bc} - p_{bc} \)

**Definition 4.2**

Contrasts of the form (4.2.1) are called main effect-or type I contrasts, contrasts of the form (4.2.2) are called interaction- or type II- contrasts.

In order to see whether a given MD \( T \) allows estimation of all type I- and type II-contrasts, or, if not, how one can generate a design \( T^* \) from \( T \) that has this property it is useful to introduce the concept of connectedness of a design. We shall simply extend the results of Srivastava and Anderson (1970).

**Definition 4.3** (Srivastava and Anderson, (1970)):

(i) Two levels in \( S \), say \( u_1 \) and \( u'_1 \), are said to be connected w.r.t. all sets \( S_a \) and \( S_{bc} \), if there exists a sequence of assemblies in \( T \), say \( t_1, t_2, \ldots, t_w \), such that
\[
\delta^{-1} E(\sum_{\rho=1}^{w} (-1)^{\rho} y_{t_\rho}) = p_{1} - p_{1}^t
\]

where \(\delta(\neq 0)\) is a constant and \(y_{t_\rho}\) denotes the response \(y\) for assembly \(t_\rho\).

(ii) If all levels of \(S_i\) are connected then \(T\) is said to be connected w.r.t. \(S_i\) (or, alternatively, w.r.t. \(F_i\)).

(iii) \(T\) is said to be type I-connected if it is connected w.r.t. all \(S_a\).

(iv) The sequence of assemblies \((t_1, \ldots, t_w)\) is said to be a chain connecting levels \(u_i\) and \(u'_i\) of set \(S_i\).

**Definition 4.4**

(i) Two pairs of level combinations in \(S_{jk}\) of the form \((u_j, u'_k)\), \((u'_j, u_k)\), with either \(u_j = u'_j, u_k \neq u'_k\) or \(u_j \neq u'_j, u_k = u'_k\), are said to be connected w.r.t. all sets \(S_a\) and \(S_{bc}\) if there exists a sequence of assemblies \(t'_1, t'_2, \ldots, t'_z\) such that

\[
\gamma^{-1} E(\sum_{\rho=1}^{z} (-1)^{\rho} y_{t'_\rho}) = \sum_{j} u_j u'_j + \sum_{k} u'_j u_k
\]

where \(\gamma(\neq 0)\) is a constant.

(ii) If all admissible quadruplets of \(S_{jk}\) as given in (i) are connected then \(T\) is said to be connected w.r.t. \(S_{jk}\) (or, alternatively, w.r.t. \(F_j \times F_k\)).

(iii) \(T\) is said to be type II-connected if it is connected w.r.t.
all $S_{bc}$.

(iv) The sequence $(t'_1, \ldots, t'_z)$ is said to be a chain connecting level combinations $(u_j, u'_k), (u'_j, u'_k)$, set $S_{jk}$. The above definitions are useful in checking the connectedness of designs.

**Definition 4.5** T is said to be completely connected if it is type I- and type II-connected. If T is not completely connected it is said to be partially connected.

To check whether a given design T has the property of type I- and type II-connectedness, the following theorems can be used. Let $\lambda^u_i(T)$ denote the frequency of level $u_i$ (Bose and Srivastava, 1964a) and $\lambda^u_{jk}(T)$ denote the frequency of level combination $(u_j, u'_k)$ in design T.

**Theorem 4.1** (Srivastava and Anderson, 1970):

The design T is type I-connected if and only if the following conditions hold: For every set $S_i$ let the sequence $(t_1, t_2, \ldots, t_{w_i})$ with $w_i$ even be such that $T_{1i} = (t_2, t_4, \ldots, t_{w_i})$ and $T_{2i} = (t_1, t_3, \ldots, t_{w_i-1})$ with

\[
\begin{align*}
(a) & \quad \lambda^u_r(T_{1i}) = \lambda^u_r(T_{2i}) \quad 1 \leq u_r \leq n_r, \ r \neq i \\
(b) & \quad \lambda^u_i(T_{1i}) = \lambda^u_i(T_{2i}) \quad v_1 \neq u_i \text{ or } u'_i \\
(c) & \quad \lambda^u_i(T_{1i}) = \lambda^u_i(T_{2i}) + \delta \\
& \quad \lambda^u_i(T_{1i}) = \lambda^u_i(T_{2i}) - \delta \\
& \quad \lambda^u_x y(T_{1i}) = \lambda^u_x y(T_{2i}) \quad 1 \leq u_x \leq n_x, \ 1 \leq u_y \leq n_y, \ all \ S_{xy}
\end{align*}
\]

where $\delta$ is a non-zero integer.
Theorem 4.1 is useful in constructing main effect designs. For designs with two-factor interaction, we have the following theorem.

Theorem 4.2

The design $T$ is type II-connected if and only if the following conditions hold: For every set $S_{jk}$ and every pair of level combinations $(u_j, u_k), (u'_j, u'_k) \in S_{jk}$ with either $u_j = u'_j$, $u_k \neq u'_k$, or $u_j \neq u'_j$, $u_k = u'_k$, let the sequence of assemblies $(t_{1jk}^i, t_{2jk}^i, \ldots, t_{zjk}^i)$ with $z_{jk}$ even be such that $T_{1jk} = (t_{2jk}^1, t_{4jk}^1, \ldots, t_{zjk}^1)$ and $T_{2jk} = (t_{1jk}^1, t_{3jk}^1, \ldots, t_{zjk}^1)$ with

(a) $\lambda_r(u_j^r)(T_{1jk}) = \lambda_r(u_k^r)(T_{2jk})$ \hspace{1cm} $1 \leq u_r < n$, $1 \leq r < m$

(b) $\lambda_x(u_j^x)(T_{1jk}) = \lambda_x(u_k^x)(T_{2jk})$ \hspace{1cm} $1 \leq u_x < n$, $1 \leq u_y < n$

(c) $\lambda_{jk}(u_j^{jk})(T_{1jk}) = \lambda_{jk}(u_k^{jk})(T_{2jk}) + \gamma$ 

$\lambda_{jk}(u_j^{k})(T_{1jk}) = \lambda_{jk}(u_k^{k})(T_{2jk}) + \gamma$

where $\gamma$ is a non-zero integer

(d) $\lambda_{jk}(v_j^{jk})(T_{1jk}) = \lambda_{jk}(v_k^{jk})(T_{2jk})$ \hspace{1cm} $(v_j, v_k) \neq (u_j, u_k)$ and

$(u_j^{k}, u_k^{k})$
Proof: Suppose the conditions (a), (b), (c) and (d) hold. Consider the expression
\[ z = \sum_{\rho=1}^{p} (-1)^{\rho} y_{\rho}. \]
By condition (a) all main effects \( p_r \), and by (b), interaction effects \( p_x y \) where \( S_{xy} \neq S_{jk} \), will each appear in this expression the same number of times with a plus sign as with a minus sign. Now consider the interaction effects of level combinations of \( S_{jk} \). All effects of level combinations of \( S_{jk} \), other than the two mentioned, \( (u_j u_k), u_j' u_k' \), will then appear in the expression with a plus and a minus sign the same number of times as indicated by (d). Then by (c), it is seen that
\[ z = \sum_{\rho=1}^{p} (-1)^{\rho} y_{\rho} = \left( p_{12} u_j u_k + u_j' u_k' \right) \]
Conversely, suppose that \( t_1', t_2', ..., t_z' \) is a chain connecting level combinations \( (u_j, u_k) \) and \( (u_j', u_k') \), of set \( S_{jk} \). It then follows directly from Definition 4.4 that conditions (a), (b), (c), and (d) must hold.

The relationship between connectedness and estimability is given by the following theorems.

Theorem 4.3 (Srivastava and Anderson, 1970).

The type I-contrast \( p_i - p_i' \) is estimable if and only if \( u_i \) and \( u_i' \) in \( S_i \) are connected w.r.t. all sets \( S_a \) and \( S_{bc} \).
Theorem 4.4

The linear function \( p_{jk} - p_{jk}' \) is estimable if and only if \((u_j, u_k)\) and \((u_j', u_k')\) in \(S_{jk}\) are connected with respect to all sets \(S_a\) and \(S_{bc}\).

Proof: If there is a chain connecting these two level combinations of \(S_{jk}\), then by Definition 4.4, the contrast is estimable. Conversely, if the contrast is estimable then there exists some linear combination of \(E(y_{t_1})\), such that

\[
\sum a_i E(y_{t_1}) = u_{jk}' - u_{jk}'
\]

where the coefficients \(a_i\) are rational numbers. To see this, we observe that since the contrast is estimable, we know that its best linear unbiased estimator is

\[
\hat{p}_{jk} - \hat{p}_{jk}'
\]

where the \(\hat{p}_{jk}\) are obtained as the solutions to the normal equations; viz. \(p_T = (X'X)^{-1}X'y\). A matrix with integer elements, \(\Gamma\), may be chosen such that a choice of \((X'X)^{-1}\) could be \((X'X + \Gamma)^{-1}\), since \((X'X + \Gamma)\) is a non-singular matrix. (See Theorem 2.3).

Knowing then that \((X'X + \Gamma)\) and \(X_T\) have integer elements, we see that the elements of \((X'X + \Gamma)^{-1}\) are rational, and have the solutions to the normal equations as linear functions of the observations with rational coefficients. The proof is then straightforward.
Following the proof of Anderson (1968), we may let $\delta$ be an integer such that $a_i = c_i \delta^{-1}$, where $c_i$ is an integer for all $i$. Then

$$\sum_{i} a_i(y_{t_i}) = \delta^{-1} \sum_{i} c_i E(y_{t_i})$$

$$= \sum_{u_ju_k} u_j' u_k' = p_{jk} - p_{jk}$$

Then we may arrange the assemblies, with $t_i$ each appearing $c_i$ times, into a sequence $(t_1, \ldots, t_z) = T^*$, such that $(t_1, t_2, \ldots, t_z) = T_1$ is formed by all assemblies with a plus sign and $(t_1, t_2, \ldots, t_{z-1}) = T_2$ is formed by all assemblies with a minus sign. Since (a), (b), (c), and (d) of theorem 4.2 must hold for this sequence $T^*$, $T^*$ is a chain connecting the level combinations $(u_j, u_k)$, $(u_j' u_k')$, of $S_{jk}$.

This completes the proof.

In order to show then that all type I and type II contrasts are estimable, we may use theorem 4.4 and its corollaries which follow directly.

**Corollary 4.4.1** Type II-contrasts are estimable if either

$$u_j u_j - p_{jk} u_j' u_j' = \sum_{u_j u_k} u_j u_k - \sum_{u_j u_k} u_j' u_k - \sum_{u_j u_k} u_j' u_k$$

and $p_{jk} - p_{jk}$ are estimable.

**Corollary 4.4.2** If the design $T$ is type II-connected, then all contrasts of the form

$$\sum_{c_{jk} p_{jk}} u_j u_k$$

with

$$\sum_{c_{jk} p_{jk}} u_j u_k = 0$$

are estimable.
Corollary 4.4.3 If the design $T$ is type I- and type II-connected, i.e. completely connected, then all type I- and type II-contrasts are estimable.

To check whether a given MD, $T$, is completely connected we use the transitivity property of connectedness and the following definition and theorem.

**Definition 4.6**

Let $T$ be a MD with $M^*(T)$ sets $S_a, S_{bc}$. Let $T_v$ be a MD design obtained from $T$ by ignoring $v$ sets, and let $S_a$ be one of the remaining sets, where $S_a$ can refer to a set of type $S_a$ or $S_{bc}$; such as $S_a$ is said to belong to $T_v$. Then $S_a$ is said to be connected w.r.t. the remaining $M^*(T) - v - 1$ sets in the original $T$ if $T_v$ is connected w.r.t. $S_a$.

**Theorem 4.5**

An MD, $T$, is completely connected if and only if there exists a sequence of designs $T_v$, $v = 0, 1, ..., M^*(T) - 2$ such that $T_v$ is connected w.r.t. to at least one $S_a$, where $S_a$ belongs to $T_v$.

**Proof:** We shall verify the first steps in this sequential procedure. Without loss of generality suppose that $S_1$ is connected with respect to every other set $S_a$ and $S_{bc}$ in the design $T_o = T$. Then for any contrast

$$
\begin{align*}
\xi & \quad \psi_1 \quad u_1 \\
\psi & \quad \psi_1 \quad p_1 \\
u & \quad \psi_1 \quad i_1
\end{align*}
$$

where the $\psi_1$ are constants with $\sum \psi_1 u_1 = 0$, there exists a linear
function of the observations, $\psi'y_T$ say, such that

$$(4.2.3) \quad E(\psi'y_T) = \sum_{i} \frac{u_i}{u_i} p_i.$$ 

Next suppose that $S_{jk}$ is connected with respect to $T$, which is formed by ignoring $S_i$ (or $F_i$) in $T$. Then according to Definition 4.2 there exists a sequence of assemblies $\sum_{i} (-1)^{\rho} t_{\rho}$ say, with $z$ even such that

$$(4.2.4) \quad E(\delta \sum_{i} (-1)^{\rho} t_{\rho} = \sum_{i} \frac{u_i}{u_i} p_i^{\prime} + p_j^{\prime} p_k - p_j k) \quad \text{for all pairs } (u_j, u_k) \text{ in } S_{jk} \text{ as specified earlier. Obviously}$$

$$\sum_{i} \frac{u_i}{u_i} p_i \text{ in (4.2.4)} \text{ is a contrast in the } p_i \text{ and hence estimable by virtue of (4.2.3). Hence } S_{jk} \text{ is connected in } T \text{ and hence by Corollary 4A.1 all type II-contrasts in } S_{jk} \text{ are estimable.}$$

Next suppose without loss of generality that $S_{\ell}$ is connected in $T_2$ which is obtained from $T$ by omitting $S_i$ and $S_{jk}$. Again, there exists a sequence of assemblies, $\sum_{i} (-1)^{\rho} t_{\rho}$ say, with $w$ even such that

$$(4.2.5) \quad E(\delta \sum_{i} (-1)^{\rho} t_{\rho} = \sum_{i} \frac{u_i}{u_i} p_i^{\prime} + \sum_{j} \frac{u_j}{u_j} p_j^{\prime} k + p_j k - p_k k) \quad \text{for all pairs } u_\ell, u_\ell' \text{ in } S_{\ell}. \text{ Obviously, since } w \text{ is even, } \sum_{i} \frac{u_i}{u_i} p_i^{\prime} \text{ is estimable because of (4.3.3) and } \sum_{j} \frac{u_j}{u_j} p_j^{\prime} k \text{ and } p_j k \text{ is estimable because of (4.3.4) and Corollary (4.4.2). Hence, } S_{\ell} \text{ is connected in } T. \text{ We}$$
proceed in this way until we reach $T_{M^*(T)-1}$.

The converse is obvious. If $T$ is completely connected, then each $T$ is connected. This completes the proof.

**Definition 4.7**

Two MD's $T$ and $T'$ are said to be model equivalent, denoted by $LM(T) = LM(T')$ if $p_T = p_{T'}$.

**Theorem 4.6**

If $T$ is a partially connected MD, then there exists a sequence of assemblies $D^* = (d_1^*, \ldots, d_n^*)$ such that $T^* = (T+D^*)$ is a completely connected design with $T$ and $T^*$ being model equivalent.

**Proof (Augmentation Procedure):** The following proof is constructive and hence provides an augmentation procedure for augmenting partially connected designs.

Let $T$ be a partially connected MD and suppose that $T$ is not connected with respect to $S_{a_1}$, $S_{a_2}$, ..., $S_{a_k}$, where $S_{a_i}$ is either of the form $S_a$ or $S_{bc}$. Let $\omega(S)$ denote the number of different factors in $S$; e.g., $\omega(S_a) = 1$, $\omega(S_{bc}) = 2$, $\omega(S_a S_{bc}) = 3$, $\omega(S_a U S_{bc}) = 3$, etc.

Then choose a set of assemblies not in $T$, say $D_1$, such that $T(1)^1 = T + D_1$ is connected with regard to $S_{a_1}$. Now ignore all elements of $S_{a_1}$ in $T(1)$ to obtain the $(m-\omega(S_{a_1}))$-dimensional design $T(1)^1 = (T+D_1)^1$.

Next choose a set of assemblies not in $T(1)^1$, say $D_2$, such that $T(1,2)^1 = T(1)^1 + D_2$ is connected with regard to $S_{a_2}$. Again, ignore all elements of $S_{a_2}$ in $T(1,2)$ and denote the resulting $(m-\omega(S_{a_1} U S_{a_2}))$-dimensional design by $T(1,2)^1$. Continue this process with $S_{a_3}, \ldots, S_{a_k}$. Notice
that the sets of assemblies $D_2, \ldots, D_k$ contain only $m_2, \ldots, m_k$ factors with $m > m_2 > \ldots > m_k$. In order to make them compatible with $T$, add to $D_i$ ($i=2, \ldots, k$) arbitrary level combinations of the remaining $m - m_i$ factors not in $D_i$. Call this set of assemblies then $D_i^*$ and let $D_i^* = \sum_{i=1}^{k} D_i^*$ with $D_1^* = D_1$. Using Theorem 4.5, this completes the proof.

The augmentation procedure is, of course, not unique; different $D_i^*$ may differ in their numbers of assemblies $h(D_i^*)$ as well as in the actual assemblies. One subclass of augmented designs that is of interest is the class of designs $T^*$ that are obtained from $T$ with the minimal number of additional assemblies.

**Definition 4.8**

The $Md, T^*$, is said to be a minimal augmented multidimensional design (MAMD) with regard to the $Md, T$, if $D^*$ has the minimum number of assemblies such that $T = T + D^*$ is completely connected.

**Theorem 4.7**

For a given $Md, T$, there exists a MAMD, $T^* = T + D^*$ with the number of different assemblies in $T^*$, equal to $h(T^*)$, where $h(T^*) = h(T) + h(D^*)$, and $h(D^*) = v(p_T) - r(X_T)$, $v(p_T)$ is the number of d.f. for $p_T$, and $r(X_T)$ is the rank of $X_T$.

**Proof:** The fact that a MAMD, $T^*$, does exist follows from Theorem 4.6.

Let the d.f. due to $S_\alpha$ (independent of the design) be $v_\alpha(p_T)$ and the d.f. available for $S_\alpha$ in $T$ be $v_\alpha(p_T)$. Then $v_\alpha(p_T) - v_\alpha(p_T)$ represents the additional d.f. required to connect $S_\alpha$. Since $S_\alpha$ is not connected
in $T$, not all assemblies required to connect $S_\alpha$ are present in $T$. Therefore, $v_\alpha(p_T) - v_{\alpha_T}(p_T)$ simply represents the number of new assemblies needed to connect $S_\alpha$. If we proceed in like manner with $k$ such sets then $v(p_T) = \sum_{i=1}^k v_i(p_T)$ and $r(X_T) = \sum_{i=1}^k v_i(X_T)$. It follows that $h(D^*) = v(p_T) - r(X_T)$.

An illustration of theorem 4.7 is found in Example 4.1.

Example 4.1

$M(T) = 3, m = 4, n = n = n = 3$. The factors $F_{1i}, i=1,...,4$, are identified by $F_1(1,2,3), F_2(I,II,III), F_3(A,B,C), F_4(\alpha,\beta,\gamma)$

The Design $T$

A $3^2$ Replicate of $3^4$ Factorial

$h(T) = 9$ Assemblies

$\begin{align*}
(1,I,A,\alpha) & \quad (1) \\
(1,II,B,\beta) & \quad (2) \\
(1,III,C,\gamma) & \quad (3) \\
(2,I,B,\beta) & \quad (4) \\
(2,II,A,\gamma) & \quad (5) \\
(2,III,C,\alpha) & \quad (6) \\
(3,I,C,\gamma) & \quad (7) \\
(3,II,B,\alpha) & \quad (8) \\
(3,III,A,\beta) & \quad (9)
\end{align*}$

The above assemblies may be represented by $t_i, i=1,...,9$. For the assumed model

$E(y) = X_{12}p_{12} + X_{13}p_{13} + X_{44}p_{44}$

the design $T$ is not connected in sets $S_{12}, S_{13}, S_{14}$. A sequence to connect $T, D^*$, may be expressed in terms of the $D_j, j=1,2,3$, that follow.
Sequence | Assemblies | Set Connected
---|---|---
D | (1, I, A, \( \beta \)), (1, I, A, \( \gamma \)) | S
D : | (2, I, A) (1, II, A) (1, III, A) | S
D : | (3, I, A) (2, III, A) (3, II, A) | S
D : | Empty | S

Therefore \( h(D^*) = 8 \) and \( h(T^*) = 17 \).

In order to see that \( T^* = T + D^* \) is connected in sets \( S_4, S_{12}, S_{13} \), and \( S_4 \), observe that if for the purpose of illustration one observation is taken per assembly and if \( (1, I, A, \beta) = \gamma \), then for set \( S_4 \):

\[
(1, I, A, \alpha) - (1, I, A, \beta) = p_4^\alpha - p_4^\beta
\]

\[
(1, I, A, \alpha) - (1, I, A, \gamma) = p_4^\alpha - p_4^\gamma
\]

Now that \( S_4 \) is connected we may ignore it for the purpose of finding assemblies to connect \( S_{12} \). We need only to find assemblies in three factors. (Any levels of \( S_4 \) may be appended to these assemblies). Then, for set \( S_{12} \):

\[
(2, I, A) + (1, II, A) - (2, II, A) - (1, I, A) =
\]

\[
= p_{12}^{^2 I} + p_{12}^{III} - p_{12}^{^2 II} - p_{12}^{II}
\]

\[
(3, I, A) + (1, II, A) - (3, II, A) - (1, I, A) =
\]

\[
= p_{12}^{^3 I} + p_{12}^{III} - p_{12}^{^3 II} - p_{12}^{II}
\]
(l,III,A) + (2,II,A) - (2,III,A) - (1,II,A) = 
\[ \frac{\hat{\beta}_{I \!II \!III}}{12} + \frac{\hat{\beta}_{I \!II}}{12} - \frac{\hat{\beta}_{I \!III}}{12} - \frac{\hat{\beta}_{I \!II}}{12} \]

By theorem 4.5, \( S \) must be connected.

The augmented design of Example 4.1 would have the minimum number of assemblies needed to estimate contrasts of the sets \( S_{12}, S_{13}, \) and \( S_4 \), since \( h(T^*) = 17 \). This need not always be the case. We may add the minimum number of assemblies to augment a design \( T \) and still have \( h(T^*) > v(p_T) + 1 \), as indicated by Example 4.2 which follows.

In Example 4.2, we shall use standard notation to represent the factor level combinations, as the notation given in Example 4.1 was given for the purpose of illustration.

Example 4.2

The Design \( T \)

(Addelman and Kempthorne, 1961)
An Orthogonal Main Effect Plan For A

\( 4 \times 3^4 \times 2 \) Factorial

\( h(T) = 25 \) Assemblies

000000 (1) 101111 (6) 202221 (11) 302221 (16) 000000 (21) 011220 (2) 112200 (7) 212001 (12) 310011 (17) 010121 (22) 022011 (3) 122020 (8) 220120 (13) 320201 (18) 021201 (23) 022101 (4) 120201 (9) 220210 (14) 321020 (19) 022021 (24) 000221 (5) 100021 (10) 201001 (15) 302100 (20) 002210 (25)
The model: \[ E(y) = X_{1} p_{1} + X_{23} p_{23} + X_{24} p_{24} + X_{55} p_{55} + X_{66} p_{66} \]

The old assemblies are \( t_{i}, \) \( i = 1, \ldots, 25. \)

<table>
<thead>
<tr>
<th>Sequence</th>
<th>New Assemblies</th>
<th>Assemblies Used To connect ( S_{\alpha} )</th>
<th>Set Connected</th>
</tr>
</thead>
<tbody>
<tr>
<td>D ( 1 )</td>
<td>111220 (26)</td>
<td>( (t, t_{2}, t_{26}, t_{27}, t_{9}, t_{18}) )</td>
<td>S 1</td>
</tr>
<tr>
<td></td>
<td>211220 (27)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D ( 2 )</td>
<td>Empty</td>
<td>( (t, t_{8}, t_{24}) )</td>
<td>S 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( (t, t_{10}, t_{3}, t_{8}) )</td>
<td>S 5</td>
</tr>
<tr>
<td>D ( 4 )</td>
<td>110211 (28)</td>
<td>( (t, t_{26}, t_{22}, t_{28}, t_{1}) )</td>
<td>S 23</td>
</tr>
<tr>
<td></td>
<td></td>
<td>( (t, t_{23}, t_{15}) )</td>
<td></td>
</tr>
<tr>
<td>D ( 5 )</td>
<td>Empty</td>
<td>Empty</td>
<td>S 24</td>
</tr>
</tbody>
</table>

\( h(D) = 3, h(T) = 28, v(p_{1}) = 20. \)

Note here that using assemblies \( t_{2} \) and \( t_{26} \) we may obtain an estimate of \( p_{1} - p_{1}. \) In like manner assemblies \( t_{27}, t_{9}, \) and \( t_{18} \) may be used to demonstrate that all type \( 1 \)-contrasts of \( S_{1} \) are estimable, and hence \( S_{1} \) is connected. "Assemblies Used To Connect \( S_{\alpha} \)" then connect \( S_{6}, \) \( S_{5}, S_{23}, \) and \( S_{24} \) in a similar manner.

**Corollary 4.7.1**

For a MAMD, \( T^{*} \), the number of assemblies in \( D^{*}, \) \( h(D^{*}) \), is independent of the ordering of the sets \( S_{\alpha}, S_{\alpha}, \ldots, S_{\alpha}. \) in the augmentation procedure.

**Proof:** Obvious from the augmentation procedure.
CHAPTER V

OPTIMALITY OF MINIMAL AND MINIMAL AUGMENTED DESIGNS

5.1 Introduction

We established in Chapter IV a procedure for constructing MAMD's. For a given design T, a sequence D* may be chosen such that h(D*) is a minimum number. In this chapter we shall present a procedure for determining which D* is best or "optimal" for augmenting T. This procedure also gives rise to a method of constructing a design, T*, such that h(T*) is a minimum number; i.e. h(T*) = \nu(p_T^*) + 1.

Note that if T* is a MAMD, h(T*) is not necessarily a minimum number. As an illustration, in Example 4.2, h(T*) = 28, and

\[ \nu(p_T^*) = 3 + 4(2) + 1 + 2*2 + 2*2 = 20. \]

Therefore another class of designs whose total number of assemblies are such that h(T*) = \nu(p_T^*) + 1 shall be investigated in this chapter.

Due to the difficulty that arises in constructing MD's subject to having h(T*) or h(D*) a minimum number, the usual optimal properties such as balance, symmetry or orthogonality may not apply. For this reason we shall be led to an approach to these designs which is limited in scope. However, this approach shall be readily applicable.

There is no previous work on MAMD's or designs which require h(T*) to be a minimum number. Addelman and Kempthorne (1961) present orthogonal main effect plans for k^n factorials, some of which require the minimum number of assemblies, k(n-1) + 1, but these designs are
very few, since their goal was the construction of orthogonal designs, not minimal designs. Raktoe and Federer (1968) give an approach to non-orthogonal main effect plans in \( k^n \) factorials. This method requires \( kn \) assemblies for \( k \) even and \( k(n-1) + 2 \) assemblies for \( k \) odd. We shall present designs for \( k \) even or odd that require \( k(n-1) + 1 \) assemblies. Also, our approach will not be limited to \( k^n \) factorials.

5.2 Summary of Chapter V

Section 5.3 covers the relationship between MAMD's and designs which require a minimum number of assemblies. Section 5.4 discusses optimality criteria that are used on main effect designs and introduces the A-optimality criterion that we will use for our minimal designs. Section 5.5 outlines a procedure for determining a set of "optimum" designs, \( \Psi \), from which one finds the A-optimal design with respect to \( \Psi \). This procedure is discussed further in section 5.6 and methods for determining a set \( \Psi \) are given in sections 5.7 and 5.8. Examples will be given in 5.7 and 5.8 in order to help the reader understand these methods. Since 5.7 and 5.8 deal specifically with main effect designs, particular examples of first order interaction designs developed from these main effect designs are given in section 5.9.

5.3 MMD's and MAMD's

Before considering optimality criteria, let us clearly understand the difference between the MAMD's defined in Chapter IV and MD's which require a minimum number of assemblies.

Definition 5.1

If \( T \) is a completely connected MD with d.f. \( \nu(p_T) \) for \( p_T \) in
h = h(T) assemblies, then T is said to be a minimal multidimensional design (MMD) if h = v(p_T) + 1.

Example 5.1

Let T be a completely connected main effect MMD of m factors.

\[ v(p_T) = \sum_{i=1}^{m} (n_i - 1) = \sum_{i=1}^{m} n_i - m, \text{ and } h = \sum_{i=1}^{m} n_i - m + 1. \]

Some MAMD's are MMD's. This is illustrated by the design \( T^* = T + D^* \) in Example 4.1, where \( h(T^*) = 17 = 16 + 1 \) and \( v(p_T) = 4(2) + 4 + 4 = 16. \) As we will see later, optimal MMD's may be found by using a method similar to the method used to find optimal MAMD's. Therefore the construction of optimal MAMD's shall be considered first (section 5.5). Before doing this we need to discuss some optimality criteria to be used on MAMD's and MMD's.

The Use Of Main Effect Criteria To Determine Optimal MAMD's and MMD's.

The criteria we shall discuss in section 5.4 are used most frequently for main effect designs and are concerned with minimizing the variance or average variance of \( p_a - \hat{p}_a. \) One of these criteria shall thus serve as a guide to some MD's where two factor interactions are assumed, since the parameter \( p_{bc} \) is handled in a manner similar to the parameter \( p_a. \) For example, in order to show that the set of factor levels, \( S_{bc} \) is connected in T, simply demonstrate that all differences, \( p_{bc} - p_{bc}, (u_b u_c) \neq (v_b v_c), \) are estimable, since all interaction contrasts of the form \( p_{bc} + \)
\[ u_b^'u_c^' - u_b^'u_c - u_b^'u_c \] can be expressed in terms of the "two-way" differences,

\[ \frac{u_b^'u_c - u_b^'u_c}{p_{bc} - p_{bc}} = (p_{bc} - p_{bc}) \]

Therefore, it is possible to use main effect optimality criteria on the parameters \( p_{bc} \) in order to compare designs where a two-factor interaction is assumed.

However, in future discussion we shall develop a general construction of designs in \( M^*(T) \) sets, \( S_\alpha, \alpha = 1, \ldots, M^*(T) \). These sets are to be understood sets of factor levels, \( S_\alpha \) or sets of level combinations, \( S_{bc} \). Generally we speak of sets \( S_\alpha \). However, we shall not consider sets of the form \( S_{bc}', S_{bd}', \) say; i.e., we shall limit our general developments to designs where if first order interaction is assumed, no factor shall interact with two or more factors. Although we give a general approach to these latter designs (see Case 3 and 4 of section 5.9), the particular procedures to be mentioned in section 5.6, 5.7, and 5.8 will not apply in general. More shall be said on these types of designs in section 5.9.

5.4 A Review of Some Main Effect Optimality Criteria

Previous work in optimal designs by Plackett and Burman (1946) Chernoff (1953), Ehrenfeld (1955), Kempthorne (1956), and Kiefer (1959a, 1959b, 1961) will be summarized here.

A design may be A-, E-, or D-optimal. An A-optimal design is
one where the average variance of differences \( \hat{p}_{a}^{u} - \hat{p}_{a}^{u'} \) is least. An E-optimal design has the variances of these differences bounded by the least upper bound. A D-optimal design estimates these differences with a minimum generalized variance. Quite often A-optimal designs are both E and D-optimal. (Kiefer, (1959b). Therefore we shall concern ourselves with A-optimality.

The following theorem gives a relationship between the latent roots of the c.m. for \( p_{\alpha} \) and A-optimality. This theorem will apply to all sets \( S_{\alpha} \), for parameter \( p_{\alpha}, \alpha = 1, \ldots, M(T) \). In the following theorem if \( p_{\alpha} = p_{bc} \), then \( n_{\alpha} = n_{b} n_{c} \).

**Theorem 5.1 (Kempthorne, 1956):**

Let \( \hat{C}_{\alpha} \) be the \( n_{\alpha} \times n_{\alpha} \) coefficient matrix of the reduced normal equations, \( \hat{C}_{\alpha} \hat{p}_{\alpha} = \hat{Q}_{\alpha} \). Further, let the rank of \( \hat{C}_{\alpha} \) be \( n_{\alpha} - 1 \), and denote by \( \lambda_{i}, i = 1, 2, \ldots, n_{\alpha} - 1 \), the \( n_{\alpha} - 1 \) positive latent roots of \( \hat{C}_{\alpha} \). Then

\[
\frac{1}{n_{\alpha} (n_{\alpha} - 1)} \sum_{\alpha, \alpha' \neq \alpha} \left( \hat{p}_{\alpha}^{u} - \hat{p}_{\alpha}^{u'} \right) \hat{v}_{\alpha}^{u} \hat{v}_{\alpha}^{u'} = \frac{2}{n_{\alpha} - 1} \sum_{i=1}^{n_{\alpha} - 1} \frac{1}{\lambda_{i}} \sigma^{2}
\]

where \( \hat{V}(\hat{p}_{\alpha}^{u} - \hat{p}_{\alpha}^{u'}) \) is the variance of \( \hat{p}_{\alpha}^{u} - \hat{p}_{\alpha}^{u'} \).

Thus, a design \( T^{*} \) will be A-optimal for set \( S_{\alpha} \) if the c.m. for \( \hat{p}_{\alpha} \) has roots \( \lambda_{i} \) such that \( \sum_{i=1}^{1} \frac{1}{\lambda_{i}} \) is a minimum.
Ordinarily, for experimental designs, one approaches the concept of optimality in terms of the variance-covariance matrix involved. Since \( C_\alpha \) is a singular we have no variance-covariance matrix in the strict sense. However, we shall relate A-optimality to the g-inverse for \( C_\alpha \). If \( C_\alpha \) is an \( n_\alpha \times n_\alpha \) c.m. for \( \hat{p}_\alpha \) of rank \( n_\alpha - 1 \) and

\[
(C_\alpha + J)^{-1}
\]

(see theorem 2.2) is used as a g-inverse for \( C_\alpha \), then

\[
\text{tr}(C_\alpha^-) = \sum_{i=1}^{n_\alpha-1} \frac{1}{\lambda_i} + \frac{1}{n_\alpha}
\]

where \( n_\alpha \) is known and independent of the design (see Hinkelmann, 1968a). Therefore, a design that is A-optimal for \( S_\alpha \), has a c.m. for \( \hat{p}_\alpha \), \( C_\alpha \), such that there exists a g-inverse, \( C_\alpha^- \), where \( \text{tr}(C_\alpha^-) \) is a minimum.

5.5 Finding Optimal MAMD's

The Method

We will indicate here how A-optimality can be applied to the problem of augmenting a partially connected design \( T \) for one set \( S_\alpha \). (The problem of augmenting \( T \) for more than one set shall be discussed later). As indicated in Chapter IV there may be many sequences, \( D^* \) that will connect \( S_\alpha \) with the minimum number of assemblies. In order to find an optimal \( D^* \), and thus an optimal \( T^* = T + D^* \), our method is as follows: 1) consider a set of possible sequences that will connect \( S_\alpha \) in \( T^* \), i.e. determine an optimum set of designs, \( \Psi \) (defined by a property to be mentioned in this section and constructed by methods of
section 5.7 and 5.8). This optimum set is to be generated either from
the given design to be augmented, $T_a^\psi$ (which we will use for MAMB's).
or a constructed design, $T_b^\psi$ (which we will use for MMD's). Next,

(2) in the set $\psi$, find the A-optimal design for set $S_\alpha$.

Before we apply this method let us first make a change in
notation. For a set $\psi$ of optimum designs, an optimal MAMD will now
be designated by $T^*$ and an optimal sequence of assemblies by $D^*$. The
members of $\psi$ will be MAMD's, and will be designated by $T_i$ with corres-
ponding sequence $D_i$, $i=1,\ldots,z$, say.

**Definition 5.2**

Let $\psi = (T_1, T_2, \ldots, T_z)$ be a set of MD's. Let $T^*$ be a member of
this set and $C_\alpha^*$ the coefficient matrix for $p_\alpha$ in $T^*$. $T^*$ is said to
be A-optimal for $p_\alpha$ with respect to $\psi$ if for any other design $T_i$ in
$\psi$ whose corresponding c.m. has roots $\lambda_j$,

$$\sum_{j=1}^{\alpha-1} \frac{1}{\lambda_j^*} < \sum_{j=1}^{\alpha-1} \frac{1}{\lambda_j}$$

where $\lambda_j^*$ are the non-zero roots of the c.m. for $\hat{p}_\alpha$ in $T^*$. Since the
only optimality criterion we are considering is A-optimality, we shall say
in what follows that $T^*$ is $\psi$-optimal for set $S_\alpha$.

**The Set $\psi$**

We next consider the problem of constructing the set $\psi$. The mem-
bers of $\psi$ are already partially defined from the design to be augmen-
ted, $T_\psi$. Each member of $\Psi$ starts with the design to be augmented, $T_\psi$, and differs only in the assemblies which are added to connect the set $S_\alpha$. The members of $\Psi$ are then $T_i = T_\psi + D_i$, where $D_i$ is constructed according to Theorem 4.5 and so that $S_\alpha$ is connected in $T_i$.

There are three ways the $D_i$ may be determined. There may be an orthogonal or balanced MAMD, $T^*_i$, that results from $T_\psi$ by adding a certain set of assemblies, $D^*_i$. Certainly we should not ignore such a possibility or such a set $D_i$.

Next, due to the experimental conditions, the factor level combinations and thus the sequences, $D_i$ may be limited to a very few. However, these two latter cases are rare and in this section we shall concern ourselves with a general approach, the third possibility, to the construction of $D_i$ and a particular set $\Psi$.

A General Approach To Finding MAMD's For $\Psi$

Consider the c.m. in the r.n.e. for $T_\alpha$,

\[(5.5.1) \quad C_\alpha = H_{\alpha\alpha} - (H_{\alpha 1} \ldots H_{\alpha m})H^{-1} (H_{\alpha 1}, \ldots, H_{\alpha m})' \]

An investigation of the composition of $C_\alpha$ may provide some information on $\Psi$-optimal MAMD's. The construction of $T_i = T_\psi + D_i$ and thus the calculation of $C_\alpha$ for each $T_i$, is limited by two things over which we have no choice. These are (1) the actual construction of $T_\psi$ and (2) the number of assemblies in $D_i$, $h(D_i)$. We shall acknowledge these two constraints when constructing $\Psi$ using a general approach.
Since $T_\alpha$ is already fixed, many elements of the matrices used to calculate $C_\alpha$ of the $T_\alpha$ in (5.5.1) are already fixed. Without loss of generality, now let $\alpha = 1$. Let us concern ourselves with the matrix $H = (H_{uv})$, $u,v = 2,\ldots,M(T)$, whose g-inverse shall be used in (5.5.1); i.e., we are ignoring how the levels of the set $S_\alpha$, the set to be connected in $T_\alpha$, appear in $T_\alpha$. Trying to describe any given $T_\alpha$ in general terms is almost impossible, and thus trying to describe the matrix $H = (H_{uv})$, $u,v = 2,\ldots,M(T)$, in general terms may be impossible. Therefore, let us direct our attention to the diagonal elements of $H$ (the level frequencies of each factor set $S_\beta$, $\beta \neq 1$).

There are cases where, if for each factor all levels of the same factor appear with equal frequency, a balanced, or symmetrical design, or one with like properties results. For example, Bose and Srivastava (1964b) indicate that this property is one requirement of an MPBD. This property, that all levels of the same set, $S_\beta$, $\beta = 2,\ldots,M(T)$, appear with equal frequency then may be viewed as a desirable characteristic. (However, this is not a rule, Addelman and Kempthorne (1961) give some examples of orthogonal designs where factor levels appear with proportional frequencies.) Further, in accordance with the definition of a good fraction given in Chapter I, this attempt at achieving equal frequencies may establish some uniformity of variance on effects involving sets $S_\beta$, $\beta = 2,\ldots,M(T)$.

Therefore, from the calculation of $C_\alpha$, we have some idea of what might be the preferred sequences, $D_1$. If possible, the level combina-
tions of all sets \( S_\beta, \beta \neq 1 \), in \( T_1 = T_\psi + D_1 \), should be such that the level (or level combinations) frequencies of each set \( S_\beta \) be equal (to \( k_\beta \) for example).

In constructing the sequences \( D_1 \), the next difficulty we face is that \( h(D_1) \) is to be a minimum. Thus it may not be possible to have the level combinations of \( S_\beta \) appear with equal frequency in \( T_1 = T_\psi + D_1 \). Let \( h(T_1) = a_\beta (\mod n_\beta) \). Then it is possible to have \( a_\beta \) levels of \( S_\beta \) occur one more time than \( n_\beta - a_\beta \) levels of \( S_\beta \). If \( h(T_1) = k_\beta n_\beta + a_\beta \), then we have

\[
(5.5.2) \quad h... = k_\beta, \quad \text{for } n_\beta - a_\beta \text{ levels of } S_\beta
\]

\[
= k_\beta + 1, \quad \text{for } a_\beta \text{ levels of } S_\beta
\]

Although it will not always be possible, we recommend that the experimenter try to establish (5.5.2) for sets \( S_\beta, \beta \neq 1 \) in constructing the designs \( T_1 \) of \( \psi \).

Federer (1961) used this type of an approach in adding new treatments to block designs. He required \( r \) treatments to appear \( v_r \) times and \( \ell \) treatments to appear \( v_\ell \) times, where \( r + \ell \) is the total number of treatments. As mentioned, due to a possibly unusual nature of \( T_\psi \), (5.5.2) may not be obtainable. Then it may be possible to have for set \( S_\beta \), \( r_j \) levels occurring \( v_{r_j} \) times, \( j=1,2,\ldots \), with \( \sum r_j = n_\beta \).
However, those cases where the frequencies of each level in $S_\beta$ are not equal in $T_\psi$ are unusual. The MD's most frequently used are those with balance, partial balance, or orthogonality that require the level frequencies of $S_\beta$ to be the same. (The exception is again given in some cases of Addelman and Kempthorne, 1961). Therefore, most $T_\psi$'s to be augmented will have equal level frequencies in each $S_\beta$.

Now all members of our set $\Psi$, $T_1 = T_\psi + D_1$, are to be the same w.r.t. level combinations of all sets $S_\beta$, $\beta \neq 1$. This means that ignoring $S_1$, all $D_1$ are to be the same. Further in determining the $D_1$ such that (5.5.2) holds for all $T_1$, we also determine the off diagonal matrices of $H, H_{uv}, u \neq v, u, v = 2, \ldots, M(T)$. In summary then, all designs in $\Psi$ will use the same $H$ matrix for calculating the c.m. C_1.

To achieve different designs $T_1$, in $\Psi$, we vary the levels of $S_1$ in combination with levels of all other sets $S_\beta$. Exactly how this can be done and still achieve designs $T_1$ that are connected in $S_1$ shall be discussed in sections 5.7 and 5.8. Let us for the moment assume that this can be done. Then all the members of $\Psi$ will differ only in the occurrence of levels of $S_1$. One of these $T_1$ will then be $\Psi$-optimal for set $S_1$. 
A Summary of Our General Approach To Constructing $\Psi$

Let us review our procedure that we have just discussed for finding a $\Psi$-optimal MAMD. One begins with a design $T_{\Psi}$, not connected in set $S_\alpha$. A set $\Psi$ of possible MAMD's is considered and an $\Psi$-optimal design for $S_\alpha$ is found.

A general approach to constructing members of $\Psi$ is follows:

1. The sequences $D_i$ are to be constructed using levels of sets $S_\beta$, $\beta \neq \alpha$, by the augmentation procedure of Theorem 4.6 and in light of (5.5.2) (Procedure I, section 5.7 will demonstrate exactly how this can be done); (2) next, the sequences, $D_i$, are to have levels of $S_\alpha$ occurring with different frequencies (Procedure II, section 5.8 will demonstrate exactly how this can be done).

An example of this method is seen in Appendix II, where a 1/3 replicate of a 3 factorial is augmented for the presence of a two-factor interaction. There, $S_\alpha = S_{12}$ and $S_\beta = S_3$. For $S$ in $T^*$

$$h.. \ell = 3 \quad \text{for } \ell = \text{level III}$$
$$= 4 \quad \text{for } \ell = \text{levels I, II}.$$  

The set $\Psi$ then consists of all possible level combinations of $S$ with $S_3$, such that (5.5.3) holds for $S$.

For More Than One Set $S_\alpha$

Now let us consider the problem of augmenting design $T_{\Psi}$ for more than one set $S_\alpha$, say $S_{\alpha_1}, S_{\alpha_2}, \ldots, S_{\alpha_k}$. The procedure mentioned in this section still applies, where from $\Psi$ we determine a design $T^*$ optimal
for one set, say $S_{\alpha_1}$. For the design $T_{\psi_{\alpha}}$, sequences $D_{\alpha}$ can be constructed such that one attempts to maintain (5.5.2) for all sets $S_{\beta}$, $\beta \neq \alpha_1$. (Note, $S_{\alpha_2}, \ldots, S_{\alpha_k}$ are included here). Then the members of $\Psi, T_{\alpha}$, will differ only in the manner that levels of $S_{\alpha_1}$ occur with the factor level combinations from the other sets $S_{\beta}$, $\beta \neq \alpha_1$.

Some reasons why we attempt to find a MAMD that is optimal for one set $S_{\alpha_1}$ shall follow. A common case of a MAMD may be a design $T_{\psi_{\alpha}}$ that is augmented for only one set $S_{\alpha}$. For example, the case where a main effect design is augmented for the presence of one two-factor interaction. Common main effect plans will not have many suspected interactions, since if many are suspected, a first-order interaction design perhaps would have initially been suggested. More reasons why we give a method of find $\Psi$-optimal MAMD's for one set are given in the next section.

5.6 Reasons For Finding MMD's and MAMD's That Are $\Psi$-optimal For Only One Set, $S_{\alpha}$

The method of finding $\Psi$-optimal MAMD's outlined in section 5.5 shall also apply to MMD's. More will be said on this method in the following sections, but first let us present some reasons for finding a design that is $\Psi$-optimal for only one set, $S_{\alpha}$.

Given a fixed number of assemblies, it may be preferred to seek a design which is "best" for one factor (or pair of factors) rather than seek a design that is "semi-best" for two or more factors. For example, it may be preferred to choose a design $T^*$ in $\psi$ that has minimum
average variance for estimate of \( p_a - p_a' \), rather than a design

\[ T^{**} \]

that has the minimum variance on all such sets \( p_a - p_a' \), \( a=1, \ldots, M(T) \), or on a set of estimable functions \( b_1, b_2, \ldots, b_k \). However, \( T^* \) should have some degree of uniform variance on effects not involving \( S_a \).

We shall take this view toward finding optimal MMD's (and MAMD's) for two reasons. The first is that with the tools at hand (previous developments), a procedure for finding an A-optimal MMD (or MAMD) is made easier. A unified approach to finding MAMD's (and as it will be seen for MMD's) is made possible through the use of the general form of the matrix \( C_{\alpha} \). The relationship of \( C_{\alpha} \) and a g-inverse \( C_{\alpha}^- \) was seen in section 5.4. Using the latent roots of \( C_{\alpha} \), \( \Psi \)-optimality was defined. In constructing \( \Psi \) using the form of \( C_{\alpha} \) as a guide, we attempted some "protection" for the sets \( S_\beta, \beta = 1, \ldots, M(T) \), \( \beta \neq \alpha \), by maintaining \((5.5.2)\). Thus using previous developments, we can establish one approach to finding these designs, and this fact is important, since MMD's and MAMD's may be difficult to construct.

The second reason that we take such a view is that in order to use a criterion as A-optimality on all estimates \( \hat{b}_1, \hat{b}_2, \ldots, \hat{b}_k \), one has to be able to describe the design in terms of the full rank model,

\[ E(y) = Ub, \quad b' = (b_1, \ldots, b_k). \]

If \( \lambda_i, i=1, \ldots, k \), are the latent roots of \((U'U)\), then the design that is A-optimal for the set of estimable functions \( b_1, \ldots, b_k \) is the design with the minimum sum, \( \Sigma_{i=1}^{k} \frac{1}{\lambda_i} \).
However, this approach requires some description of the construction of 
(U'U). Given that h(T*) or h(D*) is to be a minimum number, the 
description of (U'U) is difficult in general terms. Therefore, the 
construction of a set Ψ from which one identifies an A-optimal design 
for estimable functions b₁,...,bₖ is not apparent.

Let us next investigate some methods for constructing Ψ as indicated in section 5.5.

5.7 Procedure I

In this section we will present a method for constructing members 
of the set Ψ used to find optimal MAMD's. This method also presents a 
way of constructing a set Ψ to be used for finding optimal MMD's. In 
constructing the set Ψ for MAMD's, the starting point TΨₐ, the design 
to be augmented, is given. In constructing Ψ for MMD's a starting point 
TΨₐ has to be created.

In creating TΨₐ, we shall simultaneously demonstrate how the D₁ₐ 
of T₁ in the set Ψ for MAMD's may be constructed in levels from the 
sets Sβ, β ≠ α. Then in section 5.8, we will indicate ways that levels 
of Sα may be applied to the D₁ for MAMD's or to the design TΨₐ for MMD's.

For the sake of notation, in Procedure I (given in this section) 
and in Procedure II (given in section 5.8) we will assume we are con-
structing main effect designs; i.e. we have sets of levels (instead 
of level combinations) Sβ, β = 1,...,m, only. Without loss of gen-
erality, the construction of designs with sets Sβ, β =1,...,M*(T),
follows in the same manner, as we will demonstrate by the examples given in section 5.9. For the following definition, the reader should recall that $h_{\beta}, \ldots$ is the frequency of level $\ell_\beta$ of set $S_\beta$.

**Design Class I**

A MMD in $h$ assemblies with factors $F_1, \ldots, F_m$ with factor $F_\beta$ having $n_\beta$ levels, is said to belong to Design Class I if, for $\beta = 1, \ldots, m$,

\[
(1) \quad h = k_\beta n_\beta + a_\beta \quad \text{and} \\
(2) \quad h_{\beta}, \ldots = k_\beta \quad \text{for } n_\beta - a_\beta \text{ levels of } S_\beta \quad = k_\beta + 1 \quad \text{for } a_\beta \text{ levels of } S_\beta
\]

Procedure I, though not too difficult to apply is difficult to describe in general terms. Therefore, it may be helpful to the reader to refer to Example 5.2 and other examples as often as necessary in this and in the upcoming section.

**A Summary of Procedure I**

1. In order to construct a MMD in $m$ factors, we first construct an MMD in two factors, then an MMD in three factors, \ldots, and finally an MMD in $m$ factors.

2. In each case we use the same method:

   a) To construct a MMD of $k$ factors we "augment" a design of $k-1$ factors.

   b) Starting with a design of $k-1$ factors $F_1, F_2, \ldots, F_{k-1}$, in $h^{(k-2)}$ assemblies, $n_k - 1$ additional assemblies in $k-1$ factors are to be added to the original $h^{(k-2)}$ such that (5.7.1) holds for all $k-1$ factors.
c) These \( n_k - 1 \) additional assemblies should be repeats of
\( n_k - 1 \) assemblies of the original \( h^{(k-2)} \). There are two reasons for this

1) In doing so, (5.7.1) is easier to maintain
2) Levels of \( S_k \) can be appended to the \( h^{(k-2)} + n_k - 1 \) factors in a simple manner.

d) If c) can not be attained, the some of the \( n_k - 1 \) assemblies have to be new assemblies in \( k - 1 \) factors. These new assemblies have to be chosen so that (5.7.1) is attained and
levels of \( S_k \) can be easily connected in the design of
\( h^{(k-2)} + n_k - 1 \) assemblies of \( k \) factors.

**Procedure I For Constructing Designs of Class I**

**m=2**

We will first construct a main effect plan for a \( n_1 \times n_2 \) factorial in \( h^{(1)} \) assemblies, where \( h^{(1)} = n_1 + n_2 - 1 \). A design in levels of
\( S_1 \) and \( S_2 \) is constructed accordingly:

(a) For the purposes of describing Procedure I, we will first choose the set \( S_\ell \), \( \ell = 1 \) or 2 with the largest \( n_\ell \). Without loss of generality, let \( n_1 \geq n_2 \). Also let \( \phi_i, i=1,\ldots,n_1 \) be the
levels of set \( S_1 \). We may then list \( \phi_i \) in a column of \( h^{(1)} \) elements,
where \( h^{(1)} = a^{(1)}_1 (\mod n_1) \). For this case of two factors \( a^{(1)}_1 = n_1 - 1 \).
Note that $h_{1} = k(1) + 1$ for $a(1)$ levels of $S_{1}$

$$= k(1)$$

for $n_{1} = a(1)$, levels of $S_{1}$

with $k(1) = 1$.

(b) Next we associate levels of $S_{2}$, $\theta_{j}$, $j=1,...,n_{2}$, with the $h(1)$ levels of $S_{1}$, in such a manner that all differences $\theta_{j} - \theta_{j'}$ are estimable. This may be achieved in many ways.

We shall present an insight to some of these methods.

(i) If $n_{1} = n_{2}$, then one approach may be as follows

(5.7.3) $\begin{array}{c}
\phi_{0} \\
\phi_{0} \\
\vdots \\
\phi_{0} \\
\phi_{1} \\
\vdots \\
a_{a + 1}
\end{array}$

$h(1)$ Assemblies In Levels
Of $S_{1}$ and $S_{2}$
Note that all contrasts in levels of set $S$ are now estimable (i.e. assemblies $\phi \theta$ and $\phi \theta$ can be used to estimate $\theta - \theta$, etc.) and therefore by theorem 4.5, so are all contrasts in levels of $S$. Note also $h. j = k^{(1)}_{2} + 1$

$$= k^{(1)}_{2}$$

for $(a^{(1)}_{2} = a^{(1)}_{1})$ levels of $S$

$$= k^{(1)}_{2}$$

for $n - a^{(1)}_{2}$ levels of $S$

(And as a consequence of $m = 2$, $k^{(1)}_{2} = k^{(1)}_{1} = 1$)

(ii) If $n > n_{2}$: here we will introduce a method that one can use for later cases of $m > 2$ factors. We shall also use this method on factor level combinations in establishing Procedure II in section 5.8.

First it should be mentioned again that $h^{(1)}_{1} = n^{(1)}_{1} + n_{2} - 1$; and $h^{(1)}_{2} = n + a^{(1)}_{2}$. The levels of $S$ will be appended to levels of $S$ so as to keep $h. j = k^{(1)}_{2} + 1$

$$= k^{(1)}_{2}$$

for $a^{(1)}_{2}$ levels of $S$

$$= k^{(1)}_{2}$$

for $n - a^{(1)}_{2}$ levels of $S$

(Note, once more, that here $k^{(1)}_{2} = 1$).

One way of doing this may be as follows:
\[\phi (1) \theta (1) + 1\]

\[(5.7.4)\]

\[\phi \theta \\
\phi \theta \\
. . . \\
\phi \theta \\
\phi \theta \\
\phi \theta \\
. . . \\
\phi \theta \n + 1 \theta \omega_1 \n + 1 \theta \omega_1 \n + 1 \theta \omega_1 \n + 1 \theta \omega_1 \]

\[h(1)\] assemblies in levels of \[S_1\] and \[S_2\]

A discussion of \[\theta_1, \ldots, \theta_n - n\] is relevant here. These levels of \[S_2\] will be chosen by the experimenter so that \[h_j = k(1) + 1\] or \[k(1)\]. The last \[a(1) = n_1 - 1\] assemblies in \[S_1\] and \[S_2\] were chosen so that differences, \[0_j - 0_j\] of \[S_2\] are estimable.

Set \[S_1\] is then connected by Theorem 4.5.

In constructing designs of \(m\) factors, this last method will be used frequently. In place of \[\phi_1\] (the levels of \[S_1\]) we will be discussing \[\phi_1\] (level combinations of \(m-1\) factors). Then the levels \[0_j\] will be of set \[S_m\] and our concern will be connecting set \[S_m\].

More shall be said on this later, but notice here that since the
levels $\phi_i$, $i=1,...,a^{(1)}$ are repeated, set $S_2$ is easily connected. 

Therefore in later discussions, we would like the level combinations $\phi_i$, to be repeated, so that $S_m$ is easily connected. We shall use this idea for $m = 3$.

$\therefore m = 3$

For a minimal effect plan for a $n \times n \times n$ factorial, $h^{(2)} = n + n + n - 2 = h^{(1)} + n - 1$ assemblies are required. For three factors we proceed as if we were constructing a MMD in two factors.

For the $h^{(1)}$ assemblies of (5.7.4) let $\phi_1 = (\phi_1 \theta_1), \phi_2 = (\phi_2 \theta_2),...$,

$\phi^{(1)} = (\phi^{(1)}_1 \theta^{(1)}_1)$. We will first get $h^{(2)}$ assemblies in the first two factors. The $h^{(2)}$ assemblies in $S_1$ and $S_2$ must be such that (5.7.1) holds and levels of set $S_3$ may be easily appended and connected. Therefore we must add $n-1$ assemblies in levels of sets $S_1$ and $S_2$ to the $h^{(1)}$ assemblies such that the last statement holds.

Let these $n-1$ new assemblies in $S_1$ and $S_2$ be represented by $\phi^*_j$, $j=1,...,n-1$. As just discussed, it would be highly desirable if we could achieve (5.7.1) by merely repeating the first $n-1$ assemblies in $\phi_1$, i.e. $\phi^*_1 = \phi_1, \phi^*_2 = \phi_2, ...$, etc. Then we could connect $S_3$ as we connected $S_2$ in the $m = 2$ factor case. However, it may not always be possible that the $\phi_i$ level combinations be
repeated in all the \( n - 1 \) additional assemblies of \( S_1 \) and \( S_2 \).

(Example 5.3 indicates this for sets \( S_1 \), and \( S_2 \), and \( S_3 \), in (5.8.2).)

Therefore some of the \( \phi^*_j \) may not be equal to any of the \( \phi^*_i \), \( i=1, ... h^{(1)} \), assemblies given before. If this is the case, some adjustments have to be made in connecting \( S_3 \). (\( S_3 \) is not as easily connected as in the case where all \( \phi^*_j \)'s are repeated \( \phi^*_i \)'s.) In Table III, section 5.8, we shall give an example, of such a case for \( m=4 \). For this type of situation, one merely makes use of the \( \phi^*_j \)'s that are repeated \( \phi^*_i \)'s and thus connects some levels of \( S_3 \), and uses these connected levels of \( S_3 \) to connect levels of \( S_2 \) and \( S_1 \). More shall be said of this in Procedure II, section 5.8.

Now the \( h^{(2)} \) assemblies in \( S_1 \) and \( S_2 \) may be expressed, using \( \phi^*_j \), \( j=1, ..., n - 1 \) to represent the new assemblies in \( S_1 \) and \( S_2 \):

\[
\begin{pmatrix}
\phi_1^* \\
\phi_2^* \\
\vdots \\
\phi_{h^{(1)}}^* \\
\phi_1 \\
\phi_2 \\
\vdots \\
\phi_{n - 1}^* \\
\phi_n^* \\
\phi_{n - 1} \\
\phi_n
\end{pmatrix}
\]

(5.7.5) \( h^{(2)} \) assemblies in \( S_1 \) and \( S_2 \) only
As mentioned, $\phi_j^*$ are such that (5.7.1) holds for $S_1$ and $S_2$ and $\phi_j^*$ are chosen such that $S_3$ can be connected. Then the levels of $S_3, \alpha_1, \alpha_2, \ldots, \alpha_n$ may be appended to (5.7.5) in the following manner.

\[
\begin{array}{c|c|c}
\phi & \alpha \\
1 & 1 \\
\phi & \alpha \\
2 & 2 \\
\vdots & \vdots \\
\phi & \alpha \\
h(1) & q \\
\phi^* & \alpha \\
1 & 1 \\
\vdots & \vdots \\
\phi^* & \alpha \\
n-1 & n \\
\end{array}
\]

(5.7.6)

$h(2)$ assemblies in $S_1$, $S_2$, and $S_3$.

A discussion of $\alpha_1, \ldots, \alpha_q$ is relevant here. These levels of $S_3$ are to be chosen by the experimenter so that (5.7.1) holds for $S_3$. The choosing of levels $\alpha_1, \ldots, \alpha_q$ in conjunction with assemblies $\phi_j^*$ determines how $S_3$ is connected. (See 5.7.10 in Example 5.2 for an illustration). In this way we can construct a main effect plan for a $n_1 \times n_2 \times n_3$ factorial in $h(2) = \sum_{i=1}^{3} n_i - 2$ assemblies and still retain (5.7.1) for $S_1$, $S_2$, and $S_3$. If one proceeds in a similar manner for $m$ factors, a main effect design $h = h(m-1) = \sum_{i=1}^{m} n_i - m + 1$ assemblies can be developed.

The case for $m = 3$ is illustrated in Example 5.2.
Example 5.2

A plan of investigation requires a main effect plan for a 2 x 3 x 4 factorial in 2 + 3 + 4 - 2 = 7 assemblies.

We first establish a main effect design for a 2 x 3 factorial in $h^{(1)} = 2 + 3 - 1 = 4$ assemblies. In order to agree with (5.7.4) and for the purpose of illustration, we shall consider the plan for a 3 x 2 x 4 factorial, and consider the $h^{(1)} = 4$ assemblies of a 3 x 2 factorial first.

Let the levels of $S_1$ be 0, 1, 2, and of $S_2$ be 0, 1. Then according to (5.7.3) we have

\[(5.7.7)\]

\[
\begin{array}{c}
0 \\
1 \\
2 \\
0
\end{array}
\]

$h^{(1)} = 4$ levels of $S_1$

Then according to (5.7.4)

\[(5.7.8)\]

\[
\begin{array}{c}
00 \\
11 \\
20 \\
01
\end{array}
\]

$h^{(1)} = 4$ level combinations of $S_1$ and $S_2$

Note that if we had followed (5.7.4) exactly we would have

\[
\begin{array}{c}
00 \\
11 \\
20 \\
00
\end{array}
\]

where the (00) combination is repeated, and this is not desired, since the level 0 of set $S_2$ occurs 3 times and level 1 occurs once, which is not in accordance with (5.7.1).
Since \( h^{(1)} = 0 \) (mod 2), we find \( h_j = 2 = k_2^{(1)} \) for 2 levels of \( S_2 \), and \( h_j = 3 \) for no level of \( S_3 \); and \( h^{(1)} = 1 \) (mod 3), implies \( h_1 = 1 = k_1^{(1)} \) for 2 levels of \( S_1 \) and \( h_1 = 2 \) for 1 level of \( S_1 \).

The design in (5.7.8) is also connected since if we represent an observation on assembly (01) as \( y_{01} \), then

\[
E[y_{01} - y_{11}] = p_1 - p_1,
\]
\[
E[y_{00} - y_{01}] = p_2 - p_2,
\]
\[
E[y_{00} - y_{20}] = p_1 - p_1.
\]

For a 3 x 2 x 4 factorial, \( h^{(2)} = 7 \) assemblies are required.

Therefore to (5.7.8) we add three more assemblies as we did in (5.7.6): here \( \phi_1^{*} = \phi_1, \phi_2^{*} = \phi_2, \phi_3^{*} = \phi_3 \),

\[
\begin{align*}
00 & \quad h^{(1)} = 4 \\
11 & \\
20 & \quad h^{(2)} = 7 \text{ assemblies in } S_1 \text{ and } S_2.
\end{align*}
\]

(5.7.9)

Now append all the \( n_3 \) levels of \( S_3 \) to the first \( n_3 \) level combinations of (5.7.9) and then append \( n_3 - 1 \) levels of \( S_3 \) to the last \( n_3 - 1 \) level combinations of (5.7.9). From this we have a design connected in \( S_3 \) and thus connected in \( S_1 \) and \( S_2 \):
According to (5.7.6), $\alpha_1 = 0, \alpha_2 = 1, \alpha_3 = 2, \alpha_4 = 3, q = 4$.

Note that for (5.7.10) we have $h^{(2)} = 1 \pmod{2}$, $h^{(2)} = 1 \pmod{3}$, and $h^{(2)} = 3 \pmod{4}$, which implies $k^{(2)}_1 = 2, k^{(2)}_2 = 3, k^{(2)}_3 = 1$.

Hence

$$h_{i..} = 3 \text{ for } i = 0$$
$$= 2 \text{ for } i = 1, 2$$

$$h_{j..} = 4 \text{ for } j = 0$$
$$= 3 \text{ for } j = 1$$

$$h_{..l} = 2 \text{ for } l = 1, 2, 3$$
$$= 1 \text{ for } l = 0$$

which satisfies the characteristics of Class I designs.

5.8 Procedure II

Using Procedure I we may now construct a design $T_{\psi_b}$ (or a sequence of assemblies, $D_1$) in $m - 1$ factors. We will next consider methods to append levels of $S$ to the assemblies of $T_{\psi_b}$ (or to the assemblies of $D_1$). Without loss of generality, and to ease the presentation of our results, we shall consider appending levels of $S$ to the design...
T_b constructed in $m - 1$ factors.

One method of appending levels of $S_m$ to $T_b$, is Procedure I.

Using Procedure I, $m$ levels of $S_m$ would occur $k_m + 1$ times and $n_m - a$ levels of $S_m$ would occur $k_m$ times. Another method of appending levels of $S_m$ to $T_b$ will be Procedure II (soon to be described).

Procedure II will require $n_m - 1$ levels of $S_m$ to occur once and 1 level of $S_m$ to occur $h - n_m + 1$ times.

Given designs constructed by Procedure I and Procedure II, we may then vary the association of levels of $S_m$ in combination with the levels of the other $m - 1$ factors and in this way, construct still other members of set $\Psi$. In each case, the levels of $S_m$ will have a different occurrence and frequency of occurrence with the level combinations of the other factors.

Designs to be attained by Procedure II will be identified by Design Class II.

**Design Class II**

A MMD in $h$ assemblies with factors $F_1, F_2, \ldots, F_m$ with factor $F_\beta$ having $n_\beta$ levels is said to belong to Design Class II if for $\beta = 1, \ldots, m - 1$, the design is a member of Design Class I, and

\[
h_{\beta m} = h - n_m + 1 \quad \text{for 1 level of } S_m
\]

\[
h_{m} = 1 \quad \text{for } n_m - 1 \text{ levels of } S_m
\]
Procedure II, though not too difficult to apply, is difficult to describe in general terms. Therefore, it may help the reader to refer to Example 5.3 and other examples as often as necessary in this section.

A Summary of Procedure II

1. From a MMD of \( h - n_m + 1 \) assemblies of \( m - 1 \) factors, \( T_{\psi \theta b} \), a MMD of Design Class II in \( h \) assemblies is to be constructed.

2. To \( T_{\psi \theta b} \), add \( n_m - 1 \) assemblies in \( m - 1 \) factors such that (5.7.1) holds for all \( m - 1 \) factors in \( h \) assemblies.

3. Append one level of \( S_m \) to the first \( h - n_m + 1 \) assemblies and the remaining \( n_m - 1 \) levels of \( S_m \) to the remaining \( n_m - 1 \) assemblies.

4. The set \( S_m \) is then automatically connected in the \( h \) assemblies which form a MMD of Class II.

Procedure II For Constructing Designs of Class II

We are now going to describe Procedure II in more detail. Assume a design of \( m - 1 \) factors constructed by Procedure I in \( h^{(m-2)} = \sum_{i=1}^{m-1} n_i - m + 2 \) assemblies. Denote the \( h^{(m-2)} \) assemblies by \( \phi_i \), \( i = 1, \ldots, h^{(m-2)} \). We will need \( n_m - 1 \) additional assemblies \( \phi_j^* \)'s in \( m - 1 \) factors in order to connect \( S_m \) in \( h^{(m-1)} = h^{(m-2)} + n_m - 1 \) assemblies. As before, we would like to repeat as many \( \phi_j^* \)'s as possible in the \( \phi_j^* \)'s. If not all \( \phi_j^* \) can be repeated \( \phi_j^* \)'s , then
some of the $\phi_j^*$'s need to be constructed.

How To Construct "New" $\phi_j^*$'s

One approach is as follows: Ignore, for the moment, levels of $S_{m-1}$. It may be possible to repeat level combinations of $m-2$ factors of the $h^{(m-2)}$ assemblies in the necessary $\phi_j^*$'s so as to retain the characteristics of (5.7.1) for the $m-2$ factors. Assume that this is true. Then levels of $S_{m-1}$ have to be chosen according to the frequencies of (5.7.1) and appended to the repeated level combinations of $m-2$ factors. (See assemblies (xi) and (xii) of (5.8.2).) Thus $n_{m-1}$ additional assemblies of $m-1$ factors can be constructed according to (5.7.1).

If it is not possible to repeat level combinations of $m-2$ factors in the $\phi_j^*$'s, then we would ignore levels of sets $S_{m-2}$, $S_{m-1}$, and seek to repeat level combinations of $m-3$ factors while maintaining (5.7.1). We would then have to append according to (5.7.1) levels of sets $S_{m-2}$ and $S_{m-1}$ to the level combinations of $m-3$ factors. This method could be repeated again. Once the $\phi_j^*$'s are constructed, the next step is to append levels of $S_m$.

Connecting Levels of $S_m$

Let $0 = 1, \ldots, n_m$, be the levels of $S_m$. Let us append these levels to $\phi_1^*$, and the $n_{m-1}$ added assemblies, $\phi_j^*$, in the following way:
The reason behind this construction is obvious. The first \((m-2)\) assemblies serve as a connected design in the first \((m-1)\) factors, since the same comparisons in \(h^{(m-2)}\) assemblies can be made as before. Therefore \(S_m\) is automatically connected. Also, because of the particular way in which the \(\phi_i\) are chosen, for \(h^{(m-1)} = a_\beta \pmod{n_\beta}\), we have
\[
\begin{align*}
h..._\beta &= k_\beta + 1 & \text{for } a_\beta \text{ levels of } S_\beta \\
&= k_\beta & \text{for } n_\beta - a_\beta \text{ levels of } S_\beta
\end{align*}
\]
Notice that all that is required of the \(\phi_j^*\)'s is that they satisfy (5.7.1). In constructing the \(\phi_j^*\)'s for an \(m\)-factor design according to Procedure I, the \(\phi_j\)'s must not only retain the requirements of (5.7.1), but they must also help in connecting \(S_m\). This is easily done if all the \(\phi_j^*\)'s are repeated \(\phi_i\)'s as mentioned earlier. However, when this is not true, we could refer to the method just discussed where one ignores sets \(S_{m-1}, S_{m-2}, \ldots, S_{m-k}\), repeats assemblies of
m - 2 - 1 factors such that (5.7.1) holds, and then
(1) append levels of $S_{m-1}, S_{m-2}, ..., S_{m-k}$ each according to (5.7.1),
but (2) with some intent of using the assemblies, $f_i^*, j=1, ..., n_m - 1$,
to eventually connect $S_m$. (See Tables II and III and design $T_4$
Example 5.4). This method is not too difficult to perform, since
in "sequentially" constructing $m$ factor designs according to Procedure I, we repeat as many assemblies as possible and as often as possible.

Now that we have Procedures I and II we may use them to con-
struct a set $\Psi$ from which we identify a $\Psi$- optimal MAMD. The
following examples shall illustrate this.

Example 5.3

Consider Example 5.2. If another factor, $F_4$, at 6 levels is
required in the experiment, and if particular interest is centered
on $F_4$, a MMD may be constructed using $h = 12$ assemblies or 5 more
than in (5.7.10). We first need to construct a design of three
factors in 12 assemblies. To the 7 assemblies of (5.7.10) we add 5;

<table>
<thead>
<tr>
<th>Number</th>
<th>Assembly</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(000)</td>
</tr>
<tr>
<td>(ii)</td>
<td>(111)</td>
</tr>
<tr>
<td>(iii)</td>
<td>(202)</td>
</tr>
<tr>
<td>(iv)</td>
<td>(013)</td>
</tr>
<tr>
<td>(v)</td>
<td>(001)</td>
</tr>
<tr>
<td>(vi)</td>
<td>(112)</td>
</tr>
<tr>
<td>(vii)</td>
<td>(203)</td>
</tr>
<tr>
<td>(viii)</td>
<td>(202)</td>
</tr>
<tr>
<td>(ix)</td>
<td>(013)</td>
</tr>
<tr>
<td>(x)</td>
<td>(111)</td>
</tr>
<tr>
<td>(xi)</td>
<td>(110)</td>
</tr>
<tr>
<td>(xii)</td>
<td>(200)</td>
</tr>
</tbody>
</table>

(5.8.2)

$h(2)$ = 7 assemblies in $S_1$, $S_2$, $S_3$, and $S_4$,
$n - 1 = 5$ assemblies in $S_1$, $S_2$, and $S_3$,
$\phi_j^*$'s not repeated $\phi_1$'s.
Table I

A Main Effect Plan For a 3x2x4x6 Factorial in 12 Assemblies: T₁

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>The Assemblies</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(0 0 0 0)</td>
</tr>
<tr>
<td>(ii)</td>
<td>(1 1 1 0)</td>
</tr>
<tr>
<td>(iii)</td>
<td>(2 0 2 0)</td>
</tr>
<tr>
<td>(iv)</td>
<td>(0 1 3 0)</td>
</tr>
<tr>
<td>(v)</td>
<td>(0 0 1 0)</td>
</tr>
<tr>
<td>(vi)</td>
<td>(1 1 2 0)</td>
</tr>
<tr>
<td>(vii)</td>
<td>(2 0 3 0)</td>
</tr>
<tr>
<td>(viii)</td>
<td>(2 0 2 1)</td>
</tr>
<tr>
<td>(ix)</td>
<td>(0 1 3 2)</td>
</tr>
<tr>
<td>(x)</td>
<td>(1 1 1 3)</td>
</tr>
<tr>
<td>(xi)</td>
<td>(1 1 0 4)</td>
</tr>
<tr>
<td>(xii)</td>
<td>(2 0 0 5)</td>
</tr>
</tbody>
</table>
Assemblies (xi) and (xii) are not given in (5.7.10), but were constructed in agreement with

\begin{align*}
  h_i &= 6 \quad \text{for all } i \\
  h_{i,j} &= 4 \quad " " i \\
  h_{i,k} &= 3 \quad " " k
\end{align*}

since \( h^{(3)} = 0 \pmod{3, 2, 4} \).

Also, assemblies (xi) and (xii) do repeat (11) and (20), factor level combinations that were used to construct the 7 assemblies in (5.7.10). In repeating (11) and (20), we make the construction of other designs in \( \Psi \) easier, as we shall see in constructing \( T_4 \) in Table III of Example 5.4. To finish the illustration of Procedure II, Table I now follows by repeating level 0 of \( S_4 \).

**Example 5.4**

We shall present a set \( \Psi = [T_1, \ldots, T_6] \) in this example.

\( T \) is given in Table I. \( T_1 \) will be developed by Procedure I. \( T_2 \) will follow from \( T_1 \) by interchanging levels of \( S_4 \) in \( T_1 \). \( T_3 \) will have a unique construction. \( T_5 \) and \( T_6 \) then follow from \( T_4 \), by changing levels of \( S_4 \).

We shall use (5.8.2) which was essentially developed by Procedure I in order to illustrate the construction of \( T_2 \) and \( T_4 \). For the given level combinations of \( S_1 \) and \( S_2 \), let \( A = (00) \), \( B = (11) \), \( C = (20) \), and \( D = (01) \). Then (5.8.2) becomes
Next, in order to assign levels of $S$ to (5.8.3), we note that if we connect any two of the sets $S_{\phi}(A,B,C,D)$, $S_3(0,1,2,3)$ or $S_4(0,1,...,5)$, the third will be connected by Theorem 4.5.

$T$ will be the most difficult to construct for it is of Design Class I. From (5.8.3), consider assemblies [(i), (xii)], [(ii), (v)], and [(vii), (ix)]. To each pair, append one different level of $S_{\phi}$. All levels of the $S_{\phi}(A,B,C,D)$ are then connected. We now can ignore $S_{\phi}$. To each pair of assemblies, [(vi), (xi)], [(iii), (x)], and [(viii), (iv)], append one different level of the remaining three levels of $S_{4}$. Now $S_3$ is connected, and therefore so is $S_{\phi}$. Ignoring sets $S$, $S_{\phi}$, it then follows that sets $S_1$ and $S_2$ are now connected by (5.7.8).

$T_4$ is constructed a little differently. First, some levels of $S_{\phi}$ are connected, then these levels are used to connect $S_{\phi}$ and $S_3$. Then, by Theorem 4.5, all levels of $S_{\phi}$ will be connected. By making use of assemblies (ii), (iii) and (iv) and the repetition of $B,C,D$, in (x) (viii), and (ix) respectively, we connect levels 0,1,2, and 4 of set $S_4$, as shown in Table II and use this relationship to connect $S_{\phi}$.
Table II

The Construction of $T_4$

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>Levels 0,1,2,4 of $S_4$ Connected</th>
<th>All levels of $S_4, S_5$ Connected</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(A0 )</td>
<td>(A02)</td>
</tr>
<tr>
<td>(ii)</td>
<td>(B10)</td>
<td>(B10)</td>
</tr>
<tr>
<td>(iii)</td>
<td>(C22)</td>
<td>(C22)</td>
</tr>
<tr>
<td>(iv)</td>
<td>(D30)</td>
<td>(D30)</td>
</tr>
<tr>
<td>(v)</td>
<td>(A1 )</td>
<td>(A14)</td>
</tr>
<tr>
<td>(vi)</td>
<td>(B2 )</td>
<td>(B2 )</td>
</tr>
<tr>
<td>(vii)</td>
<td>(C3 )</td>
<td>(C32)</td>
</tr>
<tr>
<td>(viii)</td>
<td>(C20)</td>
<td>(C20)</td>
</tr>
<tr>
<td>(ix)</td>
<td>(D34)</td>
<td>(D34)</td>
</tr>
<tr>
<td>(x)</td>
<td>(B11)</td>
<td>(B11)</td>
</tr>
<tr>
<td>(xi)</td>
<td>(B0 )</td>
<td>(B0 )</td>
</tr>
<tr>
<td>(xii)</td>
<td>(C0 )</td>
<td>(C00)</td>
</tr>
</tbody>
</table>
and $S_3$.

In order to demonstrate how we connect $S_3$ and $S_4$ we now refer to Table III. The primary assemblies of Table III give the comparisons necessary to estimate the differences in levels of each set, and once a set is connected, its levels then may be ignored. The secondary assemblies are those comparisons involving the four connected levels of $S$ that support the primary assemblies. For example, assemblies (ii) and (v) estimate the difference in the levels of $A$ and $B$ of $S_4$, because the difference in levels 2 and 4 of $S_4$ is made estimable by assemblies (iv) and (ix).

Note that without appending any level of $S_4$ to (vi) and (xi), $T_4$ is connected. The levels 3 and 5 of $S_4$ may be appended now to $B_0$ and $B_2$ arbitrarily. Thus all sets $S_3$, $S_4$ and $S_4$ are connected in $T_4$. Now once again ignore levels of $S_3$ and $S_4$. Then in Table IV, sets $S_1$ and $S_2$ are connected by (5,7,8), and $T_5$ is connected.

Design $T_6$ follows from $T_5$ by interchanging levels of $S_4$ in $T_3$ in the following manner, 0$\leftrightarrow$1, 1$\leftrightarrow$3, 2$\leftrightarrow$4, 3$\leftrightarrow$2, 4$\leftrightarrow$5, 5$\leftrightarrow$0. $T_6$ follows from $T_5$ by replacing the levels of $S_4$ in (v), (i), and (vii) with 0. $T_5$ follows from $T_4$ by changing the first assembly to (0002).

Now that we have constructed a set of designs, $\Psi$, we may identify a $\Psi$-optimal design. If $\lambda_i$, $i=1,\ldots,5$ are the non-zero latent roots for each matrix $C_4$ in design $T_j$, $j=1,\ldots,6$, then we see in Table V that $T_2$ and $T_3$ are $\Psi$-optimal for $S_4$. 
Table III

Assemblies That Connect $S_0, S_3$

<table>
<thead>
<tr>
<th>Set</th>
<th>Levels Connected</th>
<th>Primary Assemblies</th>
<th>Secondary Assemblies</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S_0$</td>
<td>A, B</td>
<td>(ii), (v)</td>
<td>(iv), (ix)</td>
</tr>
<tr>
<td>$S_0$</td>
<td>A, C</td>
<td>(i), (xii)</td>
<td>(iii), (vii)</td>
</tr>
<tr>
<td>$S_0$</td>
<td>C, D</td>
<td>(vii), (ix)</td>
<td>(iii), (viii),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(iv), (ix)</td>
</tr>
<tr>
<td>$S_3$</td>
<td>0, 1</td>
<td>(i), (iii)</td>
<td>(iii), (vii)</td>
</tr>
<tr>
<td>$S_3$</td>
<td>1, 2</td>
<td>(iii), (x)</td>
<td>(ii), (x),</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(iii), (viii)</td>
</tr>
</tbody>
</table>
Table IV

A Set \( \mathbf{\Psi} \) For a 3x2x4x6 Factorial

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>( T_1 )</th>
<th>( T_2 )</th>
<th>( T_3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(0000)</td>
<td>(0005)</td>
<td>(0000)</td>
</tr>
<tr>
<td>(ii)</td>
<td>(1110)</td>
<td>(1110)</td>
<td>(1111)</td>
</tr>
<tr>
<td>(iv)</td>
<td>(0130)</td>
<td>(0132)</td>
<td>(0134)</td>
</tr>
<tr>
<td>(v)</td>
<td>(0010)</td>
<td>(0010)</td>
<td>(0011)</td>
</tr>
<tr>
<td>(vi)</td>
<td>(1120)</td>
<td>(1124)</td>
<td>(1125)</td>
</tr>
<tr>
<td>(vii)</td>
<td>(2030)</td>
<td>(2033)</td>
<td>(2032)</td>
</tr>
<tr>
<td>(ix)</td>
<td>(0132)</td>
<td>(0133)</td>
<td>(0132)</td>
</tr>
<tr>
<td>(x)</td>
<td>(1113)</td>
<td>(1111)</td>
<td>(1113)</td>
</tr>
<tr>
<td>(xi)</td>
<td>(1104)</td>
<td>(1104)</td>
<td>(1105)</td>
</tr>
</tbody>
</table>

\# with level frequencies of \( S_4 \):

<table>
<thead>
<tr>
<th>( \lambda )</th>
<th>( h \ldots \ell )</th>
<th>2</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>7</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Table IV (continued)

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>( T_4 )</th>
<th>( T_5 )</th>
<th>( T_6 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>( (0002) )</td>
<td>( (0000) )</td>
<td>( (0002) )</td>
</tr>
<tr>
<td>(ii)</td>
<td>( (1110) )</td>
<td>( (1110) )</td>
<td>( (1110) )</td>
</tr>
<tr>
<td>(iii)</td>
<td>( (2022) )</td>
<td>( (2022) )</td>
<td>( (2022) )</td>
</tr>
<tr>
<td>(iv)</td>
<td>( (0130) )</td>
<td>( (0130) )</td>
<td>( (0130) )</td>
</tr>
<tr>
<td>(v)</td>
<td>( (0014) )</td>
<td>( (0010) )</td>
<td>( (0010) )</td>
</tr>
<tr>
<td>(vi)</td>
<td>( (1125) )</td>
<td>( (1125) )</td>
<td>( (1125) )</td>
</tr>
<tr>
<td>(vii)</td>
<td>( (2032) )</td>
<td>( (2030) )</td>
<td>( (2030) )</td>
</tr>
<tr>
<td>(viii)</td>
<td>( (2020) )</td>
<td>( (2020) )</td>
<td>( (2020) )</td>
</tr>
<tr>
<td>(ix)</td>
<td>( (0134) )</td>
<td>( (0134) )</td>
<td>( (0134) )</td>
</tr>
<tr>
<td>(x)</td>
<td>( (1111) )</td>
<td>( (1111) )</td>
<td>( (1111) )</td>
</tr>
<tr>
<td>(xi)</td>
<td>( (1103) )</td>
<td>( (1103) )</td>
<td>( (1103) )</td>
</tr>
</tbody>
</table>

# with level frequencies of \( S_4 \):

<table>
<thead>
<tr>
<th>( l )</th>
<th>4</th>
<th>7</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>
Table V

The Sum, $\sum_{\lambda_1}^1$, For Members Of $\Psi$, Example 5.4

<table>
<thead>
<tr>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
<th>$T_4$</th>
<th>$T_5$</th>
<th>$T_6$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sum_{\lambda_1}^1$</td>
<td>13.34</td>
<td>12.00</td>
<td>12.00</td>
<td>18.36</td>
<td>13.34</td>
</tr>
</tbody>
</table>
We have just presented an example of how one may construct a set \( \Psi \). More examples shall be given in section 5.9 where \( \Psi \)-optimal designs with assumed two factor interactions are presented.

5.9 Examples of Minimal Designs With Two-Factor Interactions

Some MMD's with two-factor interactions present are easy to construct. The process follows the construction of main effect designs indicated in sections 5.7 and 5.8. Other MMD's with interaction do not have ease of construction, but may be constructed using some guidelines of optimality. Examples given in this section will be for different cases of \( 3^n \) factorials, since these plans are easy to illustrate and understand. However, these methods extend easily to more complex cases, i.e. mixed factorials.

Case (1): A \( 3^3 \) factorial under the model \( E(y) = X_1 p_{12} + X_3 p_{33} \)

For this plan a 9x3 main effect plan is first constructed, since there are 9 (=3x3) combinations of \( S_{12} \) and 3 levels of \( S_3 \). For this design, 11 assemblies are required. We need to first construct members of the set \( \Psi \). Then find a \( \Psi \)-optimal design, \( T^* \). This is done in Appendix II, and the \( \Psi \)-optimal design for \( S_{12} \) is given in Table VI. Once a \( \Psi \)-optimal MMD for set \( S_{12} \) is found, assemblies in sets \( S_{12}, S_1, S_2 \) and \( S_3 \) follow by the correspondence indicated in Table VI.

Coincidently, the design indicated by Table VI may also be viewed as a \( \Psi \)-optimal MAMD (as indicated in Appendix II). The first nine
### Table VI
A $\Psi$-Optimal Design For Case (1)

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>Assemblies of $S_{12}, S_2$</th>
<th>Assemblies of $S_1, S_2, S_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>$0$</td>
<td>$0 0 0$</td>
</tr>
<tr>
<td>(ii)</td>
<td>$1$</td>
<td>$0 1 0$</td>
</tr>
<tr>
<td>(iii)</td>
<td>$2$</td>
<td>$0 2 0$</td>
</tr>
<tr>
<td>(iv)</td>
<td>$3 1$</td>
<td>$1 0 1$</td>
</tr>
<tr>
<td>(v)</td>
<td>$4 1$</td>
<td>$1 1 1$</td>
</tr>
<tr>
<td>(vi)</td>
<td>$5 1$</td>
<td>$1 2 1$</td>
</tr>
<tr>
<td>(vii)</td>
<td>$6 2$</td>
<td>$2 0 2$</td>
</tr>
<tr>
<td>(viii)</td>
<td>$7 2$</td>
<td>$2 1 2$</td>
</tr>
<tr>
<td>(ix)</td>
<td>$8 2$</td>
<td>$2 2 2$</td>
</tr>
<tr>
<td>(x)</td>
<td>$1 1$</td>
<td>$0 1 1$</td>
</tr>
<tr>
<td>(xi)</td>
<td>$1 2$</td>
<td>$0 1 2$</td>
</tr>
</tbody>
</table>

# Column (b) follows from (a) using the correspondence: $S_{12} \rightarrow S_1 \times S_2: 0 \rightarrow (00), 1 \rightarrow (01), 2 \rightarrow (02), 3 \rightarrow (10), \text{etc.}$
assemblies of column (b) form a 1/3 replicate of a $3^3$ factorial that may serve as an orthogonal main effect plan. The last two assemblies then "augment" this plan for the presence of a factor one and two interaction.

Also, column (a) of Table VI was constructed (using Procedure II) so that estimates of all differences $p_{12}^{12} - p_{12}^{12}$, $(u, u) \neq (v, v)$ can be found. Although these differences may not be meaningful, they do permit one to apply Theorem 4.4. Thus set $S_{12}$ can be handled as if it were a set of levels rather than level combinations. This idea is used again in Case (2) and later in Case (5).

**Case (2):** A $3^4$ factorial under the model $E(y) = X p_{12}^{12} + X p_{34}^{34}$

For this case let contrasts of $p_{34}^{34}$ be considered "more important" than those of $p_{12}^{12}$. We will then determine a set $\Psi$ and from $\Psi$ find a design which is $\Psi$-optimal for the set $S_{34}$. Using Procedure I and Procedure II we need to find 17 assemblies of levels of $S_{12}$ and $S_{34}$, i.e. $h(1) = 9 + 9 - 1 = 17$. A set $\Psi$ for Case (2) is given by Table VII. $T_1$ is constructed using Procedure II and $T_2$, $T_3$, and $T_4$ follow from Procedure I and by varying levels of $S_{34}$.

For $T_\ell$, $\ell = 1, \ldots, 4$

$$h_i = 1, \quad i = 8$$

$$h_i = 2, \quad \text{otherwise}$$

and for $T_1$

$$h_j = 9, \quad j = 8$$
Table VII

A Set $\Psi$ For Case (2)

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
<th>$T_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>(0 0)</td>
<td>(0 0)</td>
<td>(0 0)</td>
<td>(0 0)</td>
</tr>
<tr>
<td>(ii)</td>
<td>(1 1)</td>
<td>(1 1)</td>
<td>(1 1)</td>
<td>(1 1)</td>
</tr>
<tr>
<td>(iii)</td>
<td>(2 2)</td>
<td>(2 2)</td>
<td>(2 2)</td>
<td>(2 2)</td>
</tr>
<tr>
<td>(iv)</td>
<td>(3 3)</td>
<td>(3 3)</td>
<td>(3 3)</td>
<td>(3 3)</td>
</tr>
<tr>
<td>(v)</td>
<td>(4 4)</td>
<td>(4 4)</td>
<td>(4 4)</td>
<td>(4 4)</td>
</tr>
<tr>
<td>(vi)</td>
<td>(5 5)</td>
<td>(5 5)</td>
<td>(5 5)</td>
<td>(5 5)</td>
</tr>
<tr>
<td>(vii)</td>
<td>(6 6)</td>
<td>(6 6)</td>
<td>(6 6)</td>
<td>(6 6)</td>
</tr>
<tr>
<td>(viii)</td>
<td>(7 7)</td>
<td>(7 7)</td>
<td>(7 7)</td>
<td>(7 7)</td>
</tr>
<tr>
<td>(ix)</td>
<td>(8 8)</td>
<td>(8 8)</td>
<td>(8 8)</td>
<td>(8 8)</td>
</tr>
<tr>
<td>(x)</td>
<td>(0 8)</td>
<td>(0 7)</td>
<td>(0 1)</td>
<td>(0 7)</td>
</tr>
<tr>
<td>(xi)</td>
<td>(1 8)</td>
<td>(1 8)</td>
<td>(1 2)</td>
<td>(1 8)</td>
</tr>
<tr>
<td>(xii)</td>
<td>(2 8)</td>
<td>(2 7)</td>
<td>(2 3)</td>
<td>(2 6)</td>
</tr>
<tr>
<td>(xiii)</td>
<td>(3 8)</td>
<td>(3 8)</td>
<td>(3 4)</td>
<td>(3 7)</td>
</tr>
<tr>
<td>(xiv)</td>
<td>(4 8)</td>
<td>(4 7)</td>
<td>(4 5)</td>
<td>(4 8)</td>
</tr>
<tr>
<td>(xv)</td>
<td>(5 8)</td>
<td>(5 8)</td>
<td>(5 6)</td>
<td>(5 6)</td>
</tr>
<tr>
<td>(xvi)</td>
<td>(6 8)</td>
<td>(6 7)</td>
<td>(6 7)</td>
<td>(6 7)</td>
</tr>
<tr>
<td>(xvii)</td>
<td>(7 8)</td>
<td>(7 8)</td>
<td>(7 8)</td>
<td>(7 8)</td>
</tr>
</tbody>
</table>
For $T_2$

$$h_{j} = \begin{cases} 1, & j = 7, 8 \\ 5, & \text{otherwise} \end{cases}$$

For $T_3$

$$h_{j} = \begin{cases} 1, & j = 0 \\ 2, & \text{otherwise} \end{cases}$$

and for $T_4$

$$h_{j} = \begin{cases} 4, & j = 7, 8 \\ 3, & j = 6 \\ 1, & \text{otherwise} \end{cases}$$

By using Table VIII we find that $T$ is $\Psi$-optimal for set $S_{34}$. The design $T$ of sets $S_3$, $S_1$, $S_2$, and $S_4$ follows from the correspondence given in Table IX.

Given that set $S_{12}$ is connected, we can ignore it. Then we are left with 18 assemblies of sets $S_3$ and $S_4$. Set $S_3$ is obviously connected and therefore so is set $S_4$. We may do the same for showing that sets $S_1$ and $S_2$ are connected.

**Case (3):** A $3^3$ factorial under the model $E(y) = X_{12} \ p_{12-12} + X_{13} \ p_{13-13}$

Here is a case where because of the above model, the main effect approach of Procedures I and II are not applicable. However, note that in section 3.3 the matrices, $C^i_{12}$, $i=1,...,n$, are elements of a diagonal matrix, $C_{12}$, in the r.n.e. for the parameter $p_{12}$. Therefore a g-inverse for $C_{12}$ may be found from finding a g-inverse for $C_{12}$.
Table VIII

The Sum, $\frac{1}{\lambda_i}$, For Each Member of $\Psi$, Case (2)

<table>
<thead>
<tr>
<th>$T_i$</th>
<th>$T_1$</th>
<th>$T_2$</th>
<th>$T_3$</th>
<th>$T_4$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Sigma_{\lambda_i}$</td>
<td>20.18</td>
<td>22.70</td>
<td>18.94</td>
<td>24.89</td>
</tr>
</tbody>
</table>
Table IX

The Design \( T_3 \) of Case (2)

<table>
<thead>
<tr>
<th>Assembly Number</th>
<th>Assemblies in ( S_{12}, S_{34} )</th>
<th>Assemblies in ( S_1, S_2, S_3, S_4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i)</td>
<td>........................ (0 0)</td>
<td>(0000)</td>
</tr>
<tr>
<td>(ii)</td>
<td>........................ (1 1)</td>
<td>(0101)</td>
</tr>
<tr>
<td>(iii)</td>
<td>........................ (2 2)</td>
<td>(0202)</td>
</tr>
<tr>
<td>(iv)</td>
<td>........................ (3 3)</td>
<td>(1010)</td>
</tr>
<tr>
<td>(v)</td>
<td>........................ (4 4)</td>
<td>(1111)</td>
</tr>
<tr>
<td>(vi)</td>
<td>........................ (5 5)</td>
<td>(1212)</td>
</tr>
<tr>
<td>(vii)</td>
<td>........................ (6 6)</td>
<td>(2020)</td>
</tr>
<tr>
<td>(viii)</td>
<td>........................ (7 7)</td>
<td>(2121)</td>
</tr>
<tr>
<td>(ix)</td>
<td>........................ (8 8)</td>
<td>(2222)</td>
</tr>
<tr>
<td>(x)</td>
<td>........................ (0 1)</td>
<td>(0001)</td>
</tr>
<tr>
<td>(xi)</td>
<td>........................ (1 2)</td>
<td>(0102)</td>
</tr>
<tr>
<td>(xii)</td>
<td>........................ (2 3)</td>
<td>(0210)</td>
</tr>
<tr>
<td>(xiii)</td>
<td>........................ (3 4)</td>
<td>(1011)</td>
</tr>
<tr>
<td>(xiv)</td>
<td>........................ (4 5)</td>
<td>(1112)</td>
</tr>
<tr>
<td>(xv)</td>
<td>........................ (5 6)</td>
<td>(1220)</td>
</tr>
<tr>
<td>(xvi)</td>
<td>........................ (6 7)</td>
<td>(2021)</td>
</tr>
<tr>
<td>(xvii)</td>
<td>........................ (7 8)</td>
<td>(2122)</td>
</tr>
</tbody>
</table>

# Column (b) follows from (a) using the correspondence: \( S_{12} \rightarrow S_1 \times S_2, \) \( 0 \rightarrow (00), 1 \rightarrow (01), \) etc., and \( S_{34} \rightarrow S_3 \times S_4; \) \( 0 \rightarrow (00), 1 \rightarrow (01), \) etc.
each $C_{12}^{i}$. If $\lambda_{j}^{i}$, $j = 1,\ldots,n_{2} - 1$, are the non-zero latent roots for each $C_{12}^{i}$, $i = 1,\ldots,n_{1}$, then

$$\sum_{i=1}^{n_{1}} \sum_{j=1}^{n_{2}-1} \frac{1}{\lambda_{j}^{i}}$$

can be used to find a $\Psi$-optimal design from a given set $\Psi$.

The constructing of a set $\Psi$ of minimal designs similar to Case (3) may be very difficult. Because of this difficulty, the actual construction of a minimal design, may be in itself an optimal criterion. We shall present a particular example of Case (3) and demonstrate by this example how one could generate a set $\Psi$.

In order to construct an MMD for Case (3) fifteen assemblies are required, i.e. $h = 3(2) + 2(2)(2) + 1 = 15$. Eleven assemblies can be chosen that lead to four independent contrasts in the $p_{12}^{ij}$ and four independent contrasts in the $p_{13}^{ik}$. Once these assemblies are constructed, four more assemblies can be constructed such that all contrasts of one set are estimable. Then by Theorem 4.5, the contrasts of the other set are also estimable.

To show this, let $(i,j,k)$ denote the observation and the assembly. For the purposes of illustration, let $i = 1,2,3$, and $j = A,B,C$, and $k = I,II,III$, and let $p_{12}^{ij} = ij$ and $p_{13}^{ik} = ik$. Then 11 assemblies (summarized in Table X) can be constructed such that
\[ E[(1AI) + (2BII) - (1BII) - (2AI)] = [1A + 2B - 1B - 2A] + [1I + 2II - 1II - 2I] \]

\[ E[(3AII) + (2CIII) - (3GIII) - 2AII)] = [3A + 2G - 3G - 2A] + [3II + 2III - 3III - 2II] \]

\[ E[(2CIII) + (3BI) - 2BI) - (3CIII)] = [2C + 3B - 2B - 3C] + [2III + 3I - 2I - 3III] \]

\[ E[(1CIII) + (2BII) - (1BII) - (2CIII)] = [1C + 2B - 2C - 1B] + [1III + 2II - 1II - 2III] \]

where \( E[] \) denotes the expected value.

The design in Table X is connected by the addition of (2BI), (1CII), (2CII), and (3CII), since

\[ E[(1AI) + (2BII) - (1BII) - (2AI)] = 1A + 2B - 1B - 2A \]
\[ E[(3AII) + (2CII) - (3CII) - (2AII)] = 3A + 2C - 3C - 2A \]
\[ E[(2CII) + (3BII) - (2BII) - (3CII)] = 3B + 2C - 3C - 2B \]
\[ E[(1CII) + (2BII) - (2CII) - (1BII)] = 1C + 2B - 2C - 1B \]

The fifteen assemblies are now summarized in Table XI.

A set \( \Psi \) now can be constructed by using different interaction contrasts (other than those given) to be estimated. Then from such a set \( \Psi \), an A-optimal design for \( S \) can be found.

An extension to the model of Case (3) where three interactions are present will be given next.

Case (4) A 3 factorial under the model \( E(y) = X_{12} p_{12} + X_{13} p_{13} + X_{23} p_{23} \). This case is similar to Case (3) in that the main effect approach of Procedure I and II is not applicable. We shall present a particular
Table X

11 Assemblies of The Required 15 For Case (3)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I</td>
<td>II</td>
<td>III</td>
</tr>
<tr>
<td>2</td>
<td>I, II</td>
<td>I, II</td>
<td>III</td>
</tr>
<tr>
<td>3</td>
<td>II</td>
<td>I</td>
<td>III</td>
</tr>
</tbody>
</table>
Table XI

One Design of 15 Assemblies For Case (4)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I</td>
<td>I, II</td>
<td>II, III</td>
</tr>
<tr>
<td>2</td>
<td>I, II</td>
<td>I, II</td>
<td>II, III</td>
</tr>
<tr>
<td>3</td>
<td>II</td>
<td>I</td>
<td>II, III</td>
</tr>
</tbody>
</table>
example of Case (4), and a set of similar designs \( \Psi \) may then be constructed as in Case (3), and a \( \Psi \)-optimal design for \( S_{12} \), say, can then be identified.

For Case 4, nineteen assemblies are required with all contrasts in \( S_{12} \), \( S_{13} \), and \( S_{23} \) being estimable. We shall use the notation of Case (3). To the fifteen assemblies given in Case (3) we add (1AII), (1AIII), 1BIII), and (1CI) so that contrasts in \( S_{23} \) are estimable. Let \( p_{23}^{jk} = jk \). Then

\[
\begin{align*}
E[(\text{LAI})] + (\text{LBD}) - (\text{LAI}) - (\text{LBI})] &= (AII + BII - AII - BII) \\
E[(\text{LBI})] + (\text{LCI}) - (\text{LCI}) - (\text{LBI})] &= (BII + CII - CI - BII) \\
E[(\text{LAI})] + (\text{LCI}) - (\text{LCI}) - (\text{LAI})] &= (AI + CII - CI - AII) \\
E[(\text{LAI})] + (\text{LCII}) - (\text{LCII}) - (\text{LAI})] &= (AI + CII - CI - AII)
\end{align*}
\]

These assemblies are summarized in Table XII.

By using different contrasts in \( S_{12} \), \( S_{13} \), and \( S_{23} \) to be estimated, one may construct other minimal designs in these sets and compare them on the basis of \( \Psi \)-optimality for set \( S_{12} \).

In summary then for models of the type

\[
\begin{align*}
\text{Case (1)} & \quad E(y) = X_{12} p_{12} + X_{13} p_{13} \\
\text{Case (2)} & \quad E(y) = X_{12} p_{12} + X_{34} p_{34}
\end{align*}
\]

our procedures for constructing a set \( \Psi \) may be easily applied. For Case (1) we were able to vary all levels of \( S_{12} \) in order to obtain \( \Psi \). In Case (2) we varied some levels of \( S_{34} \) and identified the \( \Psi \)-opti-
TABLE XII
One Design Of 19 Assemblies For Case (4)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>I,II,III</td>
<td>I,II,III</td>
<td>III,II,1</td>
</tr>
<tr>
<td>2</td>
<td>I,II</td>
<td>I,II</td>
<td>III,II</td>
</tr>
<tr>
<td>3</td>
<td>II</td>
<td>I</td>
<td>III,II</td>
</tr>
</tbody>
</table>
mal design. However, for cases (3) and (4)

Case (3): \[ E(y) = X_{12-12} p_{12} + X_{13-13} p_{13} \]

Case (4): \[ E(y) = X_{12-12} p_{12} + X_{13-13} p_{13} + X_{23-23} p_{23} \]

we were not able to define a set \( \Psi \) by varying levels of one set, \( S_{12} \), for example. \( \Psi \) must be defined by those contrasts one is able to estimate. For Case (3), one finds a set of assemblies involving two groups of independent interaction contrasts, then "augments" this set to estimate one group. The other group is then estimable. To achieve \( \Psi \) for Case (4), again groups of independent contrasts have to be estimated. A member of set \( \Psi \) for Case (4) was found by "augmenting" Case (3) for the presence of one more two-factor interaction. Other members of \( \Psi \) may be constructed using different contrasts and then an A-optimal member for one set chosen.
CHAPTER VI

PROPOSED EXTENSIONS

Since no previous literature is available that gives specific direction to finding designs requiring a minimum number of assemblies, it is hoped that this work has provided some grounds for future research. In this chapter we shall briefly mention some avenues of investigation.

Our approach to the analysis of incomplete designs was used because it does simplify the work involved. However, a weakness of this approach arises in analysis of designs where one factor interacts with two or more factors. This weakness appears once again in the construction of MAMD's and MMD's (as in Cases (3) and (4), Chapter V). Another problem that appears in the construction of the set from which we determine a $\Psi$-optimal design. The construction of $\Psi$ is within a limited framework, and a particular design $T^*$ is designated optimal with respect to other particular designs. Further, the design $T^*$ is optimal for only one set, $S_{\alpha}$.

Due to the extreme limitations (i.e., a minimum number of assemblies) one finds working with the general case, the difficulties just mentioned arise. Our approach was intended to remain as general as possible, applicable to symmetric and asymmetric factorials. However, solutions to the above problems may be found through the investigation of certain classes of MMD's and MAMD's. It may be possible to eliminate these restrictions (but then restrict application) by considering $k_1$ factorials, for example.
By using a full rank model, $E(y) = Ub$, the problems of analysis
would not appear. The construction of $U'U$ for $2^k$ factorials may
give some insight into the construction of a set $\Psi$, if one has to
limit oneself to such a set. The latent roots of $U'U$, $\theta_1$ could
then be related to the sum $\sum_{i=1}^{k_1} \frac{1}{\theta_i}$, the $\text{tr}(U'U)^{-1}$ and an $A$-optimal
design for all estimable functions could then be identified.

Similar developments for $3^k$ factorials may follow. From this
work, optimum plans for $2^k 3^m$ factorials may result. For these
types of special cases, and using the matrix $U'U$ as a guide, all ef-
fects will be considered in terms of optimality. Then for very
large experiments involving $2^k 3^m$ factorials, plans could result
requiring the very minimum number of necessary assemblies. For $k_1 = 30$ and $k_2 = 10$ (i.e. 40 factors), only 51 assemblies would be neces-
sary for a main effect MMD.

The next problem one may consider is the relationship of the
composition of $C_{\alpha}$, the c.m. in the r.n.e. for $p_{\alpha}$, to its latent
roots. One approach is to investigate $\text{tr}(C_{\alpha}) = \sum_{i=1}^{k_1} \lambda_i$.

A design with maximum $\text{tr}(C_{\alpha})$ permits maximum power on tests of $p_{\alpha}$. With

$$C_{\alpha} = \mathbb{I}_{\alpha\alpha} - (H_{\alpha1}, \ldots, H_{\alpha m}) (H_{\alpha1}, \ldots, H_{\alpha m})'$$

and with $\text{tr}(H_{\alpha\alpha}) = h$, the number of assemblies in $T$, $\text{tr}(C_{\alpha})$ depends
upon the diagonal elements of $[(H_{\alpha1}, \ldots, H_{\alpha m}) (H_{\alpha1}, \ldots, H_{\alpha m})']$. 
These diagonal elements are functions of $H^-$ which indirectly by Theorem 2.1 can be expressed as functions of the diagonal matrices of $H$. A further investigation may reveal how the level frequencies of sets $S_\beta$, $\beta \neq \alpha$, effect $\text{tr}(C_\alpha)$. 
CHAPTER VII

SUMMARY

In this work we considered the design and analyses of special types of \( m \)-factor designs, MMD's and MAMD's, that use a minimum number of experimental units. For cases where due to the number of required assemblies or due to the construction of these assemblies, available fractional factorials do not apply, MMD's or MAMD's may be of use. This is particularly true if there is an extreme limitation on experimental units and many factors, and thus many effects, are to be examined. However because of the nature of these designs, a previous estimate of error may be required. For this case, Cochran and Cox (1957) comment:

> If the use of an estimate of error from previous experiments is contemplated, some preliminary questions should be answered. Were the previous experiments conducted on similar experimental material, under similar conditions of experimentation and with treatment effects of about the same order of magnitude? Were the previous error variances homogeneous from one experiment to another, as indicated by a test of homogeneity of variances? If these questions are answered in the affirmative, the risk in taking a external estimate of error is decreased.

Also, another source for an estimate of error may be other experimental units not required by the design but available to the experimenter.

Due to the limitation of factor level combinations, the analysis of these designs requires some discussion. For this analysis we used a reduced form of the normal equations and thereby generalized the results of Potthoff (1958). This analysis is
applicable to all incomplete designs where two-factor interactions may or may not be assumed.

Once a form of analysis for MMD's and MAMD's is established, we proceeded to investigate methods for constructing them. An investigation into the concept of connected MD's where no interaction is assumed (Srivastava and Anderson, 1970) led us to the development of a similar concept for MD's where first order interaction is assumed. Using this latter concept we proceeded to develop a method of adding the minimum of necessary assemblies to a given design in order to estimate effects previously not estimable. In this way we developed a procedure for constructing MAMD's. This augmentation procedure enables one, for example, to construct first order interaction designs from main effect plans.

Given such a procedure, the question then arose on how to construct optimal MAMD's. Because of the limitation on the number of assemblies to be added to a design, optimality for MAMD's is difficult to define. However, using our approach to the analysis of incomplete designs we were able to present one method of finding preferred MAMD's.

This method involved the construction of a set of "optimum" MAMD's, $\Psi$. This set may be limited to a few MAMD's by the experimental situation or may be an almost infinite set. If no balanced or orthogonal MAMD's are available, we limit our investigation to a special class of MAMD's. Using this class we are then able to construct some designs for a particular set, $\Psi$, from which we identify one design which is $A$-optimal for one factor.
Design optimality is further discussed for the construction of MMD's which require the total number of assemblies used to be a minimum. Again, because of the constraint on the number used, optimality for MMD's is difficult to define. However, one approach to finding preferred MMD's similar to that of finding preferred MAMD's can be given. A set \( \Psi \) of MMD's can be generated using the procedures developed, and an A-optimal design chosen.

Using this approach to constructing main effect MMD's, examples of \( 3^n \) factorials with first order interaction were given. One of these examples is both a MMD and a MAMD, and the analysis of data generated for this example is given in Appendix I. The optimality of this particular design is discussed in Appendix II and its efficiency is compared to a more common fraction of a \( 3^3 \) factorial.

Discussions on optimality and efficiency bring us to part of the philosophy upon which this dissertation is based. An optimal design or efficient design is usually a preferred design according to a criterion that is similar to the criterion a consumer uses in buying a product. This criterion may be expressed accordingly, "Am I getting the most return for my investment?" Let us say that the cost of a product that a consumer is buying varies with size or quantity (the analogy here is with experimental units). As in the case of the more common fraction of the \( 3^3 \) factorial given in Appendix II, there may be a better investment or "bargain" in "buying" a larger size or quantity. However, a greater quantity requires a greater cost, and if this expense cannot be met by the consumer, the true bargain may be in what he can afford.
As mentioned in Chapter I, many experimental plans do not apply when the number of experimental units available to the experimenter are extremely limited or when certain factor level combinations are not available. For these cases we draw upon a well known expression dealing with affairs of the heart, and comment that it is better to have experimented once than not to have experimented at all.


APPENDIX I

A 1/3 REPLICATE OF A 3 FACTORIAL AUGMENTED
FOR AN ASSUMED TWO-FACTOR INTERACTION:
THE ANALYSIS OF DATA

Presented in Appendix I will be an application of the analysis of Model (2) given in section 3.2. For the preliminary test on fruit juices introduced in section 1.1, we shall assume a 1/3 replicate of a $\text{3}^3$ factorial is available for a test design. However, the fruit juice salesmen have indicated that a price-juice interaction may be present. In order to allow for this interaction, the nine assemblies of the 1/3 replicate given by (1AI), (2BI), (3CI), (1CII), (2AII), (3BII), (1BII), (2CIII), and (3AIII), have been augmented by two assemblies, (1BII) and (1BI). For an estimate of error, it is assumed that these eleven assemblies are to be replicated. Thus twenty-two homogeneous test periods are to be used, and it is assumed that the error involved in the application of these assemblies is normally distributed. An observation on a test period will represent the total number of 3-oz. juice cans sold during that period. The model used for this design will be

$$E(y_{ijkw}) = \mu + p_{1}^{i} + p_{2}^{j} + p_{3}^{k} + t_{12}^{ij} = p_{12}^{ij} + p_{3}^{k*}$$

where $p_{12}^{ij} = p_{1}^{i} + p_{2}^{j} + t_{12}^{ij}$ and $p_{3}^{k*} = p_{3}^{k} + \mu$. (For the r.n.e. for $p_{3}^{k}, \mu$ is associated with $p_{12}^{ij}$, i.e. $p_{12}^{ij} = p_{12}^{ij*} + \mu$)

We shall assume, for the purpose of illustration, that the experi-
ment has been performed. The data generated for this experiment shall come from values given for the parameters and error terms. We shall assume \( \mu = 500 \) and for the juices,

\[
\begin{align*}
A^1 & = 30 \\
B^1 & = 50 \\
C^1 & = -80
\end{align*}
\]

for the prices,

\[
\begin{align*}
p^1_1 & = 50 \\
p^2_2 & = -30 \\
p^3_3 & = -20
\end{align*}
\]

for the days,

\[
\begin{align*}
p^1_2 & = 40 \\
p^2_2 & = -20 \\
p^3_3 & = -20
\end{align*}
\]

and for the price-juice interaction,

\[
\begin{align*}
t^{ij}_{12} & \\
  j & = 1 \quad 2 \quad 3 \\
      i & = A \quad -20 \quad 20 \quad 0 \\
      & B \quad -10 \quad -40 \quad 50 \\
      & C \quad 30 \quad 20 \quad -50
\end{align*}
\]

For the data, error terms have been generated from a normal distribution with mean 0 and variance 9. Since the data is to be in integer form, these error terms are given in the form of integers (decimals have been rounded to whole numbers.) The error terms are then for the first replication, respectively

\[
(0,4,5,3,-2,-1,-3,2,-1,1,2),
\]

and for the second replication, respectively

\[
(-4,-2,0,5,-3,1,-4,-1,2,0,1).
\]
With these parameters and error terms we have the following data in Table XIII:

**TABLE XIII**

*Sales Data From Preliminary Test On Frozen Fruit Juice*

<table>
<thead>
<tr>
<th>Price-Juice-Day Level combination</th>
<th>Total Sales</th>
<th>Observation 1</th>
<th>Observation 2</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1AI)</td>
<td></td>
<td>600</td>
<td>596</td>
<td>1196</td>
</tr>
<tr>
<td>(2BI)</td>
<td></td>
<td>524</td>
<td>518</td>
<td>1042</td>
</tr>
<tr>
<td>(3CI)</td>
<td></td>
<td>395</td>
<td>390</td>
<td>785</td>
</tr>
<tr>
<td>(1CII)</td>
<td></td>
<td>483</td>
<td>485</td>
<td>968</td>
</tr>
<tr>
<td>(2AII)</td>
<td></td>
<td>498</td>
<td>497</td>
<td>995</td>
</tr>
<tr>
<td>(3BII)</td>
<td></td>
<td>559</td>
<td>561</td>
<td>1120</td>
</tr>
<tr>
<td>(1BIII)</td>
<td></td>
<td>567</td>
<td>566</td>
<td>1133</td>
</tr>
<tr>
<td>(2CIII)</td>
<td></td>
<td>392</td>
<td>389</td>
<td>781</td>
</tr>
<tr>
<td>(3AIII)</td>
<td></td>
<td>489</td>
<td>492</td>
<td>981</td>
</tr>
<tr>
<td>(1BI)</td>
<td></td>
<td>571</td>
<td>570</td>
<td>1141</td>
</tr>
<tr>
<td>(1BI)</td>
<td></td>
<td>632</td>
<td>631</td>
<td>1263</td>
</tr>
</tbody>
</table>

* Data in units of 3-oz, cans

Source: Data generated from given parameters

Following the methods of section 3.2, the r.n.e. for $p_{12}$ are found, $C_{12} \hat{p}_{12} = Q_{12}$, where $C_{12}$ and $Q_{12}$ are given by
\[ C_{12} = \begin{bmatrix}
1.50 & -0.50 & 0.00 & 0.00 & -0.50 & 0.00 & 0.00 & -0.50 \\
-0.50 & 4.33 & -0.50 & -0.50 & -0.67 & -0.67 & -0.50 & -0.50 \\
0.00 & -0.50 & 1.50 & 0.50 & 0.00 & 0.00 & 0.00 & -0.50 & 0.00 \\
0.00 & -0.50 & -0.50 & 1.50 & 0.00 & 0.00 & 0.00 & -0.50 & 0.00 \\
-0.50 & -0.50 & 0.00 & 0.00 & 1.50 & 0.00 & 0.00 & 0.00 & -0.50 \\
0.00 & -0.67 & 0.00 & 0.00 & 0.00 & 1.33 & -0.67 & 0.00 & 0.00 \\
0.00 & -0.67 & 0.00 & 0.00 & 0.00 & -0.67 & 1.33 & 0.00 & 0.00 \\
0.00 & -0.50 & -0.50 & -0.50 & 0.00 & 0.00 & 0.00 & 1.50 & 0.00 \\
-0.50 & -0.50 & 0.00 & 0.00 & -0.50 & 0.00 & 0.00 & 0.00 & 1.50 
\end{bmatrix} \]

\[ Q'_{12} = (124.5, 444.5, -88.0, -61.0, -29.5, -184.0, 16.0, 64.0, -286.5) \]

A g-inverse, \( C^- \), is then calculated and \( \hat{p} = C^{-1} Q \) follows:

\[ C^{-1}_{12} = \begin{bmatrix}
0.753 & -0.025 & -0.247 & -0.247 & 0.253 & -0.191 & -0.191 & -0.247 & 0.253 \\
-0.025 & 0.198 & -0.025 & -0.025 & -0.025 & 0.031 & 0.031 & -0.025 & -0.025 \\
-0.247 & -0.025 & 0.753 & 0.253 & -0.247 & -0.191 & -0.191 & 0.253 & -0.247 \\
-0.247 & -0.025 & 0.253 & 0.753 & -0.247 & -0.191 & -0.191 & 0.253 & -0.247 \\
0.253 & -0.025 & -0.247 & -0.247 & 0.753 & -0.191 & -0.191 & -0.247 & 0.253 \\
-0.191 & 0.031 & -0.191 & -0.191 & -0.191 & 0.864 & 0.364 & -0.191 & -0.191 \\
-0.191 & 0.031 & -0.191 & -0.191 & -0.191 & 0.364 & 0.864 & -0.191 & -0.191 \\
-0.247 & -0.025 & 0.253 & 0.253 & -0.247 & -0.191 & -0.191 & 0.753 & -0.247 \\
0.253 & -0.025 & -0.247 & -0.247 & 0.253 & -0.191 & -0.191 & -0.247 & 0.753 
\end{bmatrix} \]

For notation purposes, let \( p_{1A}^{-1} = 1A \). Then
\( \hat{p} = (1A, 1B, 1C, 2A, 2B, 2C, 3A, 3B, 3C) \)

The values of \( \hat{p}^{ij} \) are given below:

\[
\begin{array}{ccc}
  i & j & 1 & 2 & 3 \\
  A & 55.9 & 16.4 & 13.5 \\
  B & 89.4 & -21.0 & 79.0 \\
  C & 2.9 & -86.5 & -149.6
\end{array}
\]

Now, in order to define contrasts in main effects and interactions, we refer to the "f" parameters defined in (3.1.3) and (3.1.5)

\[
f_i = ^{pi}_{12} - ^{p..}_{12}
\]

\[
f_j = ^{p.j}_{12} - ^{p..}_{12}
\]

and

\[
f_{ij} = ^{p_{ij}}_{12} - ^{p_{i..}}_{12} - ^{p_{.j}}_{12} + ^{p_{..}}_{12}
\]

If we applied (3.1.4) and \( ^{p_{..}}_{12} = 0 \), we have

\[
\hat{f}_i = ^{\hat{p}_i}_{12} = ^{\hat{p}_i}_{12} = \frac{1}{3} \sum_j ^{\hat{p}_{ij}}_{12}
\]

\[
\hat{f}_j = ^{\hat{p}_{.j}}_{12} = ^{\hat{p}_{.j}}_{12} = \frac{1}{3} \sum_i ^{\hat{p}_{ij}}_{12}
\]

and

\[
\hat{f}_{ij} = ^{\hat{p}_{ij}}_{12} - ^{\hat{p}_{i..}}_{12} - ^{\hat{p}_{.j}}_{12} + ^{\hat{p}_{..}}_{12}
\]

Note that \( 1A + 2B - 1B - 2A \) is estimated by \( ^{\hat{p}_{1A}}_{12} + ^{\hat{p}_{2B}}_{12} - ^{\hat{p}_{1B}}_{12} - ^{\hat{p}_{2A}}_{12} \)

Thus all contrasts on main effects and interactions can be estimated by contrasts on \( \hat{p} \). For the following contrasts, \( \hat{f}_i = l_i \hat{p} \), the
corresponding variance of \( \ell_1 \), \( V(\ell_1) \) is given by

\[
L_1' c^{-1} L_1 \sigma^2
\]

\( V(\ell_1) \)

For prices

\[
\ell_1 = p^1 - p^3
\]

\[
\begin{array}{c|c|c}
\text{1} & \text{1} & \text{1} \\
68.5 & 0.33 \sigma^2 & \end{array}
\]

\[
\ell_2 = p^1 - 2p^2 + p^3
\]

\[
\begin{array}{c|c|c}
\text{3} & \text{1} & \text{1} \\
91.2 & 1.00 \sigma^2 & \end{array}
\]

For juices

\[
\ell_3 = \hat{p}^A - \hat{p}^C
\]

\[
\begin{array}{c|c|c}
\text{3} & \text{?} & \text{?} \\
106.3 & 0.22 \sigma^2 & \end{array}
\]

\[
\ell_4 = \hat{p}^A - 2\hat{p}^B + \hat{p}^C
\]

\[
\begin{array}{c|c|c}
\text{4} & \text{?} & \text{?} \\
-147.3 & 0.89 \sigma^2 & \end{array}
\]

For price-juice interaction, using our adopted notation,

\[
\ell_5 = 1A + 2B - 1B - 2A = -71.0
\]

\[
\begin{array}{c|c|c}
\text{5} & \text{?} & \text{?} \\
4.0 & 4.0 & \sigma^2 \end{array}
\]

\[
\ell_6 = 3A + 2C - 2A - 3C = 60.0
\]

\[
\begin{array}{c|c|c}
\text{6} & \text{?} & \text{?} \\
5.0 & 5.0 & \sigma^2 \end{array}
\]

\[
\ell_7 = 2C + 3B - 2B - 3C = 163.0
\]

\[
\begin{array}{c|c|c}
\text{7} & \text{?} & \text{?} \\
5.0 & 5.0 & \sigma^2 \end{array}
\]

\[
\ell_8 = 1C + 2B - 1B - 2C = -21.0
\]

\[
\begin{array}{c|c|c}
\text{8} & \text{?} & \text{?} \\
3.0 & 3.0 & \sigma^2 \end{array}
\]

The s.s. due to each contrast is obtained by

\[
(L_1' \hat{p} \Sigma L_1)^{-1} L_1' \Sigma L_1^{-1} (L_1' \hat{p} \Sigma L_1)
\]

and the s.s. due to main effects and interactions found in the following ANOVA:
Due to $p_{ij}^{12}$

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>$s.s.$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Due to $p_{ij}^{12}$</td>
<td>8</td>
<td>110,123.50</td>
</tr>
<tr>
<td>Due to $p_{j}^{1}$</td>
<td>2</td>
<td>22,387.71</td>
</tr>
<tr>
<td>Due to $p_{j}^{2}$</td>
<td>2</td>
<td>55,630.10</td>
</tr>
<tr>
<td>Due to $t_{ij}^{12}$</td>
<td>4</td>
<td>6,697.67</td>
</tr>
<tr>
<td>Due to $p_{k*}^{3}$</td>
<td>3</td>
<td>5,923,334.00 $^+$</td>
</tr>
<tr>
<td>Due to error</td>
<td>11</td>
<td>53.50</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>6,033,511.00</td>
</tr>
</tbody>
</table>

$^+$ uncorrected $s.s.$ (referred to, as "unadjusted $s.s.$")

The unadjusted $s.s.$ here can be received by $Y_{3}^{*} H_{33}^{-1} Y_{3}^{*}$ or, as given, by subtraction. As one would expect from the parameters used to generate the data, all main effects and interactions are significant at the 0.01 level.

For tests on the differences in days (I, II, III), we find the r.n.e. for $p_{3}^{*}$, $C_{3}^{-1} p_{3}^{*} = Q_{3}$ and calculate a g-inverse, $C_{3}^{-}$ to find

$$p_{3}^{*}$$

$$C_{3}^{-} = \begin{bmatrix} 1.333 & -0.667 & -0.667 \\ -0.667 & 1.333 & -0.667 \\ -0.667 & -0.667 & 1.333 \end{bmatrix}$$
\[
Q_3 = \begin{pmatrix}
84.0 & -38.0 & -46.0 \\
0.50 & 0.0 & 0.0 \\
0.0 & 0.50 & 0.0 \\
0.0 & 0.0 & 0.50 \\
\end{pmatrix}
\]

\[
c_3 = \begin{pmatrix}
0.0 \\
0.50 \\
0.0 \\
0.0 \\
\end{pmatrix}
\]

\[
\hat{p}_3 = (\hat{p}_1^3, \hat{p}^2_3, \hat{p}^3_3) = (42.0, -19.0, -23.0)
\]

Then for contrasts in \( p^k_3 \) (on the main effects of days) we have

\[
V(\hat{\ell}_i) = 65.0, \quad \sigma^2 = 1.0
\]

\[
\ell_9 = \hat{p}^1_3 - \hat{p}^2_3 = 57.0, \quad \sigma^2 = 3.0
\]

Again, the s.s. due to these contrasts is found and summarized:

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>s.s.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Days (due to ( p^k_3 ))</td>
<td>2</td>
<td>5,308.15</td>
</tr>
<tr>
<td>Due to ( p_{ij}^* ) ( \frac{12}{12} )</td>
<td>9</td>
<td>6,028,149.35( ^\dagger )</td>
</tr>
<tr>
<td>Error</td>
<td>11</td>
<td>53.50</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>6,033,511.00</td>
</tr>
</tbody>
</table>

\( ^\dagger \) uncorrected s.s. (referred to as "unadjusted s.s.")

The unadjusted s.s. due to \( p_{ij}^* \) can be found in this case either by
or, as given by subtraction. As expected, due to the input parameters, the main effects of days are significant at the 0.01 level.

"The presentation of results depends on their nature and the audience for whom they are intended," (Cochran and Cox, 1950) and since the purpose here is to illustrate the analysis of section 3.2, we could terminate our presentation here. However, more information on the main effects may be requested. For example, the mean responses of price 1 may be desired. To obtain an estimate of this mean response, \( p_1 + \mu \), a value for must be found. We may obtain this value from our work; however, as mentioned in section 3.4, "a weakness" of our approach is that in finding a g-inverse for \( C_{12} \), we find a variance-covariance matrix for contrasts involving \( p_{ij} \) only (and not \( \mu \), for example). The variance of \( p_1 + \mu \) would require a variance-covariance matrix for all parameters.

A value for \( \mu \) is obtained from the adjusted n.e.,

\[
X'X\hat{p}^* = X'y - X'X\hat{p}_{12},
\]

\[
Y^* H^{-1} Y^* \]
where

\[
\begin{bmatrix}
X'X_{33} &=& \begin{bmatrix} 8 & 0 & 0 \\ 0 & 8 & 0 \\ 0 & 0 & 6 \end{bmatrix} \\
X'y &=& \begin{bmatrix} 4 & 2 & 8 & 6 \\ 4 & 2 & 2 & 4 \\ 2 & 8 & 9 & 5 \end{bmatrix} \\
X'X_{3(12)} p_{12} &=& \begin{bmatrix} 50.4 \\ -375.6 \\ -32.7 \end{bmatrix} \\
\end{bmatrix}
\]

and

\[
\begin{bmatrix}
X'y - X'X_{3(12)} p_{12} &=& \begin{bmatrix} 4336.4 \\ 3848.4 \\ 2862.3 \end{bmatrix} \\
\end{bmatrix}
\]

From this equation we obtain \( \hat{p}^* \), and a value for \( \mu \) is

\( \mu = p^*_3 - p^*_3 \). From the above we have \( \hat{p}^*_3 = 542.05 \), and from previous work we have \( p^*_3 = 42.0 \). Therefore \( \mu = 500.05 \). Then

for prices

\[
\begin{align*}
\ell_{11} &= \hat{p}^*_1 + \hat{\mu} = 549.49 \\
\ell_{12} &= \hat{p}^*_2 + \hat{\mu} = 469.67 \\
\ell_{13} &= \hat{p}^*_3 + \hat{\mu} = 481.00
\end{align*}
\]

for juices

\[
\begin{align*}
\ell_{14} &= \hat{p}^*_4 + \hat{\mu} = 528.66
\end{align*}
\]
\[ \ell_{15} = p^B + \mu = 549.16 \]

\[ \ell_{16} = p^C + \mu = 422.33 \]

for days

\[ \ell_{17} = p^I + \mu = 542.05 \]

\[ \ell_{18} = p^II + \mu = 481.05 \]

\[ \ell_{19} = p^III + \mu = 477.05 \]
APPENDIX II

A 1/3 REPLICATE OF A 3 FACTORIAL AUGMENTED
FOR AN ASSUMED TWO-FACTOR INTERACTION:
OPTIMALITY

For Case 1, Section 5.9, and Appendix I, the design mentioned shall now be investigated for its optimality. The assemblies of a 1/3 replicate can be given as the following Latin square, with days (I, II, III) juices (A, B, C) and prices (1, 2, 3) in nine assemblies:

A 1/3 Replicate of a $3^3$ Factorial in 9 Assemblies: $T_a$

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>A</td>
<td>B</td>
<td>C</td>
</tr>
<tr>
<td>II</td>
<td>C</td>
<td>A</td>
<td>B</td>
</tr>
<tr>
<td>III</td>
<td>B</td>
<td>C</td>
<td>A</td>
</tr>
</tbody>
</table>

If we let the factors day (D), juices (J) and prices (P) be represented symbolically, the given design has the following alias structure obtained from the defining contrast $I = DJP^2$:

- $D = D^2JP^2 = JP^2$
- $J = DJ^2P^2 = DP^2$
- $P = DJ = DJP$
- $DP = DJ^2 = JP$

In order to estimate main effects, the assumption of no interaction is made. However, given a two-factor interaction between $J$ and $P$, the main effects are no longer estimable and the set of days $S_3$, and the juice-price level combinations, $S_{12}$, are not connected in $T_a$. Since there are 2 d.f. to be associated with set $S_3$ and 8 d.f. to be associated with $S_{12}$, and, there are only 8 d.f. available for estimating main effects and
interactions in $T_\Psi$, at least 2 or more assemblies are needed to connect $S_{12}$ (and thus $S_3$). We shall assume that interaction contrasts are of particular interest. Thus we are confronted with constructing a MAMD that is optimal for $S_{12}$. We shall determine a set $\Psi$ and then choose a design $T^*$ that is $\Psi$-optimal for $S_{12}$.

The set $\Psi$ will be determined by MAMD's, $T_i = T_{\Psi a} + D_i$, and thus the sequences $D_i$. The sequences $D_i$, shall be constructed according to (5.5.2) for set $S_3$, such that in $T_i$

$h_{\Psi} = 3$, $\xi = \text{III}$

$h = 4$, $\xi = \text{I,II}$

Here $h(T_i) = 11 = 3 \cdot 3 + 2$ or $h(T_i) = 2 \pmod{3}$ and $k_3 = 3$, $a_3 = 2$. In order to augment $T_{\Psi a}$ with 2 assemblies only two different types of $D_i$ are possible: one where a level combination of $S_{12}$ is repeated and one where such a combination is not repeated. For example, let $D_1 = [ (1\text{BI}), (1\text{BII}) ]$ and $D_2 = [ (2\text{CI}), (3\text{AIII}) ]$ be the two possible sequences used to augment $T_{\Psi a}$. Let $y_{1\text{AI}}$ be an observation for assembly $(1\text{AI})$. Then for $T_1$,

$E[y_{1\text{BI}} - y_{1\text{BII}}] = p_3^I - p_3^\text{III}$

$E[y_{1\text{BII}} - y_{1\text{BII}}] = p_3^{II} - p_3^\text{III}$

and for $T_2$,

$E[y_{3\text{AI}} - y_{3\text{AIII}}] = p_3^I - p_3^\text{III}$

$E[y_{2\text{CII}} - y_{2\text{CIII}}] = p_3^{II} - p_3^\text{III}$

Therefore the sets $S_{12}$ and $S_3$ are connected in $T_1$ and $T_2$. A c.m., $C_{12}$, is calculated for $T_1$ and for $T_2$ and the corresponding sum, $\frac{1}{A_4}$, is 12.67.
for $T_1$ and 14.67 for $T_2$. Therefore $T_1^* = T_\Psi + D_1$ is $\Psi$-optimal for $S_{12}$:

The Augmented 1/3 Replicate in 11 Assemblies: $T_1^*$

$$
\begin{array}{ccc}
1 & 2 & 3 \\
I & A,B & B & C \\
II & C,B & A & B \\
III & B & C & A \\
\end{array}
$$

Thus, in augmenting any 1/3 replicate of a $3^3$ for a set $S_{12}$, repeating one level combination of $S_{12}$, will produce a $\Psi$-optimal MAMD (and MMD) for $S_{12}$.

Having found a design that is $\Psi$-optimal for $S_{12}$, we shall now investigate the variances of contrasts of the price-juice interaction. We shall compare $T_1^*$, a 11/27 replicate of a $3^3$ factorial, with a 2/3 replicate to determine an efficiency factor for $T_1^*$.

A 2/3 Replicate of a $3^3$ Factorial in 18 Assemblies: $T_{2/3}$

$$
\begin{array}{ccc}
1 & 2 & 3 \\
I & A,C & B,C & A,B \\
II & B,C & A,B & A,C \\
III & A,B & A,C & B,C \\
\end{array}
$$

Let us consider the nine different interaction contrasts involving four level combinations of $J$ and $P$. (These 9 are all such possible contrasts.) For purpose of presentation let $p_{12} = ^1J$. Then the nine estimates, $\ell_1$, and their variances $V(\ell_1)$ in $T_1^*$, given one observation per assembly, are:
The variance of $l_1$ for the 2/3 replicate is 2.33 $\sigma^2$.

Then for an efficiency factor we have

$$E.F.(1) = \frac{\text{Average variance of 9 contrasts in } T_{2/3}}{\text{Average variance of 9 contrasts in } T_1} \cdot \frac{18}{11} = \frac{2.33}{8.67} \frac{\sigma^2}{\sigma^2} \cdot \frac{18}{11} = 0.440$$

Another efficiency factor can be given (Hoke, 1971) as

$$E.F.(2) = \frac{\text{tr} (C_{12}^-) \text{ for } T_{2/3}}{\text{tr} (C_{12}^-) \text{ for } T_1} \cdot \frac{18}{11} = \frac{4.41}{12.67} \frac{18}{11} = 0.570$$

In calculating $E.F.(2)$ one compares the average variance of

$$\hat{u} u_{1/2}^a \hat{u}^\alpha \hat{u}_{12}^\beta - \hat{v} v_{1/2}^\alpha \hat{v}_{12}^\beta, \ (u \ u_{a/2}) \neq (v \ v_{a/2}) \text{ for } T_1 \text{ with respect to } T_{2/3}.$$  

(For $\text{tr} (C_{12}^-)$, the sum $\sum \frac{1}{\lambda_i}$, $\lambda_i$ a root of $C_{12}$, is used.)

From $E.F.(1)$ and $E.F.(2)$ it may be argued that the efficiency of
T1* is somewhat low, but there are other desirable properties of T1* to consider. First, only 11 different assemblies are required for T1* as opposed to 18 for T2/3.

Given 17 experimental units, T2/3 could not be performed, but T1* could be, with 6 d.f. available for error. Further, the 17 - 11 = 6 assemblies used for an estimate of error do not have to be different, i.e. one assembly could be replicated 7 times.

Next, given the 11 assemblies, one could provide a test for interaction using the method of Tukey (1949, 1955) where the 4 d.f. for interaction are separated into 3 d.f. for interaction and 1 d.f. for error. Should a test on interaction prove significant, one would then have a design to estimate the interaction contrasts.

Finally, by recalling the C3 matrix given in Appendix I, it is seen that T1* remains orthogonal for days (I,II,III) and that the variance of p3 contrasts are essentially the same for T1* and T2/3. (The E.F. (1) for contrasts in p3 is 1.)
The vita has been removed from the scanned document
MINIMAL MULTIDIMENSIONAL DESIGNS

By
John T. Sennetti

(ABSTRACT)

Given is an investigation of experimental plans that require a minimum number of assemblies, minimal multidimensional designs (MMD's) and minimal augmented multidimensional designs (MAMD's). Such designs are incomplete factorials which allow for only some two-factor interactions. An analysis of incomplete factorials with and without interactions is developed from a reduced form of the normal equations.

\[ C_i p_i = Q_i \]

where \( p_i \) is a vector of estimates of factor \( i \) effects and \( C_i \) is the coefficient matrix for \( p_i \) and \( Q_i \) is a vector of transformed observations. The general forms for \( C_i \) and \( Q_i \) are presented.

The construction of MMD's and MAMD's is made possible from results obtained on connected designs. A definition of a connected design where two-factor interactions are assumed leads to a procedure for "connecting" experimental plans. This procedure provides a way of adding assemblies to a design in order to estimate contrasts not originally estimable in the design. Using this augmentation procedure and the minimum number of assemblies to be added, MAMD's may be constructed. MMD's follow by sequentially augmenting with the minimum number of assemblies, \( m \)-factor designs, \( m = 1,2,\ldots,m^* \), whose total number of factor level combinations are a minimum. A method for finding MMD's and MAMD's which are optimal for one factor or for a set of two-factors.
is then presented as well as some examples of MMD's and MAMD's with and without two-factor interactions. Data is generated and analyzed for a particular design which is both a MAMD and a MMD, and a discussion of this design's optimality is also given.