A RESPONSE SURFACE APPROACH TO THE MIXTURE PROBLEM WHEN THE MIXTURE COMPONENTS ARE CATEGORIZED.

by

John Andrew Cornell

Thesis submitted to the Graduate Faculty of the

Virginia Polytechnic Institute

in partial fulfillment for the degree of

DOCTOR OF PHILOSOPHY

in

STATISTICS

APPROVED:

I. J. Good, Chairman
Department of Statistics

Boyd Harshbarger
Department of Statistics

Klaus Hinkelmann
Department of Statistics

Richard G. Krutchkoff
Department of Statistics

Raymond H. Myers
Department of Statistics

Paul H. King
Dept. of Sanitary Engineering

Date

December 1968

Blacksburg, Virginia

TABLE OF CONTENTS

Chapter Page

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to Professor I. J. Good for his invaluable guidance in the preparation of this dissertation. Gratitude is also extended to Dr. Boyd Harshbarger, Chairman of the Department of Statistics, for his assistance and encouragements. The author also expresses his gratitude to Dr. Raymond H. Myers for his suggestions while reading the manuscript.

Appreciation is extended to Dr. Hubert M. Hill, Consultant, Tennessee Eastman Company, Kingsport, Tennessee for suggesting the topic of this research and providing the example problem in Chapter 8.

This research was carried out through fellowship funds available from the Department of Health, Education and Welfare on grant 5T1ES 33-05. For this support, the author is indebted.

LIST OF TABLES.

LIST OF FIGURES.

I. INTRODUCTION.

A typical problem in chemical experimentation is the determination of the optimal or best combination of the variables (materials) which are used in producing some measurable response. In an attempt to determine the best combinations of the variables, often one resorts to trial and error. Other attempts resemble "scattergun" procedures where a large number of combinations of the variables are tried. These procedures sometimes require large expenditures in terms of time and cost of experimentation and in most cases, better methods can be employed.

When a "multiple variable", that is, a combination of several variables is to be sought, it is possible to proceed by direct experimentation using familiar methods such as those developed by Box and Wilson in [7]. These methods are commonly referred to as response surface methods and are particularly appropriate when at least something (however little) is known about the effects of the individual variables. Because these methods will be used throughout the context, a review of the basic philosophy of response surface methodology follows.

1.1 Response surface methodology.

Suppose the emphasis is on studying the characteristics of some functional relationship

$$\eta = \phi(x_1, x_2, \ldots, x_k) , \qquad (1.1.1)$$

where η represents a measurable response which is dependent on the levels of the k variables $x_1, x_2, \ldots, x_k$ of some system. The function (1·1·1) can be defined as describing an unknown response surface over the system of the k variables. In the situation where little or no knowledge of the form of the function (1·1·1) is available, the use of response surface methods will help us to find those conditions or levels of the variables which in some sense give the "best" value of η.

Generally when using response surface methods, the experimenter will be concerned not with a single response but with a number of responses. For example, the experimenter will often be seeking conditions which maximize a major response η_1 such as yield while maintaining some auxiliary response η_2 such as purity at the best level possible. In this paper, because of the nature of the problem which we discuss in the next section, we shall limit the discussion to a single response η.

Frequently the first step in exploring the function (1·1·1) is to assume that (1·1·1) can be adequately approximated by a polynomial model. The polynomial model is thought to relate the observed response to the k variables in the form

$$y_u = \eta_u + \epsilon_u \ .$$

Of course the value of the observed response will never exactly equal the true value of the response at a setting

$(x_{u1}, x_{u2}, \ldots, x_{uk})$ and thus the quantity ϵ_u in the above model represents the error made in observing η_u at the point $(x_{u1}, x_{u2}, \ldots, x_{uk})$.

When we state that the relationship is in the form of a polynomial, we mean that $\eta(.)$ can be represented by the lower order terms of a Taylor Series expansion. This assumption about the model containing only lower order terms is usually valid particularly when the interest is limited to a small region. For example, if $\eta(\underset{\sim}{x}_u)$ is the response at the point $(x_{u1}, x_{u2}, \ldots, x_{uk})$ in a small region located within the factor space of the k variables, then $\eta(\underset{\sim}{x}_u)$ can be expressed by

$$\eta(\underset{\sim}{x}_u) = \gamma_0 + \sum_{i=1}^{k} \gamma_i x_{ui} + \sum_{i \leq j}^{k} \sum_{j}^{k} \gamma_{ij} x_{ui} x_{uj} + \sum_{i \leq j}^{k} \sum_{j}^{k} \gamma_{iij} x_{ui}^2 x_{uj} + \ldots , \qquad (1 \cdot 1 \cdot 2)$$

where the series ends with terms of degree d. This form (1·1·2) of the model is called a <u>model of degree d</u>.

When attempting to employ response surface methods, many aspects can influence the strategy of the experimenter. According to Box and Wilson [7], the following aspects are of utmost importance;

(i) the magnitude of the experimental error;

(ii) the complexity of the response surface;

(iii) whether or not experiments may be conducted sequentially so that each set of experiments may be designed using the knowledge gained from the previous sets of experiments. These three conditions will serve as guidelines for the development of the topics covered in this paper. These guidelines will be used in an attempt to employ response surface methods to a certain type of experimentation which we now discuss. This experimentation which is predominantly chemical in nature has been previously referred to in [10, 12, 16, 17, 20, 22] as the "mixture" problem.

1.2 Experiments involving mixtures.

A mixture experiment is defined as an experiment in which the response is a function only of the proportions of the components present in the mixture and not a function of the total amount of the mixture. The general purpose of mixture experimentation is to make possible estimates of the properties of an entire multi-component system from only a limited number of observations. These observations are taken at preselected combinations of the components (points in the experimental region), in an attempt to determine the best combinations of the components.

Unlike the usual response surface problem where the concomitant variables represent quantitative amounts, in the mixture problem the components represent proportions of a mixture or composition. These proportions must be non-negative and if expressed as fractions of the mixture, they must

sum to unity. For example, suppose there are k components in the system under study. If we represent the proportion of the i^{th} component in the mixture by x_i, then

$$x_i \geq 0 \qquad (1 \leq i \leq k), \tag{1·2·1}$$

and

$$\sum_{i=1}^{k} x_i = 1. \tag{1·2·2}$$

Since the proportion x_i could be unity, a mixture may be considered as being comprised entirely of one component. Hence the factor space containing the components may be geometrically represented by the interior and boundaries (vertices, edges, faces, etcetera) of a regular simplex. The vertices will represent mixtures consisting of single components and interior points will be the result of combining all the components.

In 1958, Scheffé [20] presented the first published work on mixture experimentation which dealt mainly with polynomial models and designs. The basis of this work was the choice of a symmetrical arrangement of design points in the simplex factor space and the fitting of carefully chosen polynomial models which have exactly the same number of terms as there are points in the associated designs. These designs are referred to as simplex-lattice designs and are characterized by their simplicity and intelligibility. Since then, additional work by Draper and Lawrence [10], Gorman and

Hinman [12] and others [15, 16, 17, 22], has appeared dealing mainly with the problem of designs and the plotting of variance contours.

Lambrakis [16] discusses a generalization of the simplex-lattice designs when the components (referred to as "major" components) themselves are mixtures of one or more minor components. Since each "major" component is a mixture, each "major" component may be geometrically represented by a simplex. The generalizations of the lattice designs are called multiple-lattice designs and are the result of multiplying together each simplex which defines the factor space of the individual "major" components. The construction of the multiple-lattice designs will be discussed in Chapter 2.

Recently, Thompson and Myers [22] expanded the design possibilities of Scheffé's mixture problem by defining a point of maximum interest centered in a region of interest in the simplex factor space and showed how the average mean square error of the predictor $\hat{y}$ could be minimized in conjunction with the application of rotatable designs. Since this concept of a region of interest is the framework upon which the most recent response surface techniques have been based, a similar development will be undertaken in this paper but with additional constraints placed on the components.

1.3 Categorizing the mixture components: Defining the problem.

In a chemical process, for example, production of Polyethelene Terephthalate, the response to be studied is the result of a chemical reaction of an acid with a glycol. In certain cases, the chemist might like to modify this composition by adding other acids and/or glycols. When only a modifying acid or glycol is added, there is no problem in the experimental design. However, when a modifying acid is added requiring also the addition of a modifying glycol, the normal mixture designs are unsatisfactory. This is partly because the mole percent of the acids in the mixture must sum to .5 (fifty percent) and the same restriction is placed on the glycols.

We are thus confronted not with the usual (Scheffé) definition of the mixture problem where the proportion of each component ranges from 0 to 1, but rather with mixtures in which the proportions are restricted by upper bounds. Also, unlike the "major" components in [16], the categories in our problem (acids constitute one category, glycols another category) are not considered as mixtures in themselves. Rather, each category contributes a fixed proportion to each mixture and every category must be present in every mixture. This is necessary in order to form a reaction or provide a measurable response.

Let us assume that the categories have equal weight in

defining or formulating the mixtures. For example, with three categories of components, each category contributes one third (percentagewise) to the total mixture. As in Section 1.1, let us represent the proportion of the i^{th} component present in the mixture by x_i. Then in addition to the constraints (1·2·1) and (1·2·2) placed on the components, the proportions of the components will also be restricted by

$$0 \leq x_i \leq \frac{1}{q} \qquad (n_{j-1}+1 \leq i \leq n_j ; 1 \leq j \leq q) \qquad (1\cdot3\cdot1)$$

and

$$\sum_{i=n_{j-1}+1}^{n_j} x_i = \frac{1}{q} \qquad (1\cdot3\cdot2)$$

where q represents the number of categories and n_j $(1 \leq j \leq q)$ is the number of components in the j^{th} category.

We have stated that the categories have equal weight in formulating the mixtures, that is, the proportion contributed by each category is $\frac{1}{q}$. This fixes the shape of the factor space of the components to handle the most general situation. For if the weights assigned to the categories are not equal, the general shape (defined in terms of the number of extremities) of the factor space which will be a convex polytope will remain as in the equal weight case with the exception of an area or volume change brought about by the difference in weights. This might best be seen by the following example.

To form the mixtures, suppose there are two categories

of components. Let x_1 and x_2 come from category 1 and x_3 and x_4 come from category 2. To form a valid mixture (where valid refers to satisfying the constraints (1.3.1), (1.3.2) and (1.2.2)), the components x_1 and x_2 cannot be present themselves without being in some combination with the components x_3 and x_4. Let us represent a mixture by some combination of the components where for example $x_i x_j$ is a mixture in which the i^{th} $(i=1,2)$ component comprises fifty percent of the mixture and the j^{th} $(j=3,4)$ component the remaining fifty percent of the mixture. Then a valid mixture might be represented by any of the following combinations:

$$\begin{aligned}
&x_i x_j && (i=1,2;\ j=3,4)\\
&x_i x_{i'} x_j && (i=1,2;\ i'=1,2;\ i'\neq i;\ j=3,4)\\
&x_i x_j x_{j'} && (i=1,2;\ j=3,4;\ j'=3,4;\ j'\neq j)\\
&x_i x_{i'} x_j x_{j'} && (i=1,2;\ i'=1,2;\ i'\neq i;\ j=3,4;\ j'=3,4;\ j'\neq j)
\end{aligned} \qquad (1.3.3)$$

where x_i in the notation implies that the i^{th} component is present $(x_i>0)$ in the mixture so that

$$x_1 + x_2 = \frac{1}{2}\,, \quad x_3 + x_4 = \frac{1}{2}\,. \qquad (1.3.4)$$

The factor space defined by (1.3.4) can be represented by the plane ABCD in Figure 1 on the next page. If in (1.3.4), we let

$$x_1 + x_2 = \frac{3}{4}\,, \quad x_3 + x_4 = \frac{1}{4}\,,$$

then the factor space is represented in Figure 1 by the plane

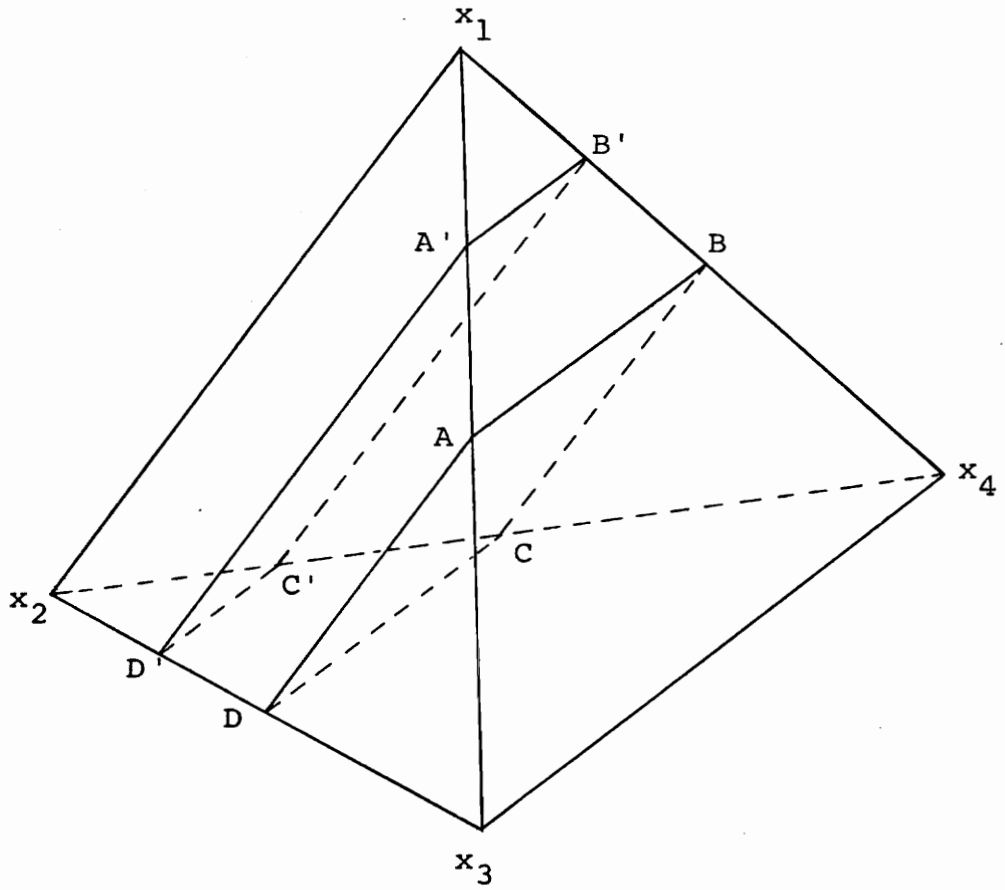

Figure 1.

Factor Space of the Mixture System in Two Dimensions.

A'B'C'D'. It is interesting to note that in defining the factor space as a regular simplex, the condition $x_1+x_2+x_3+x_4=1$ reduces four-dimensional Cartesian coordinates to "areal" (or "volume") coordinates.

Another point to mention is that in our mixture problem, each category will contain at least two components. This is because if a category consists of just one component, the problem can be reduced by scaling the proportions of the components to q-1 categories and k-1 components. Up to now (with the exception of [16]), the mixture problem has been defined as one category containing k components. It could also be defined as a two-category problem where category 1 has only one component and the other category, category 2 has k'=k-1 components so that in the second category

$$\sum_{i=1}^{k'} x_i = K < 1 \ . \qquad (1\cdot3\cdot5)$$

If we let

$$z_i = \frac{x_i}{K} \qquad (i=1,2,\ldots,k'), \qquad (1\cdot3\cdot6)$$

then

$$\sum_{i=1}^{k'} z_i = 1 \ , \qquad (1\cdot3\cdot7)$$

and thus the description is now identical with (1·2·2).

Turning our attention to the dimensionality of the problem, we know that in the usual mixture problem (one

category containing k components), because of the constraint (1.2.2), the number of independent components is k-1. In our problem however, by the additional constraint (1.3.2) placed on the components, the number of independent components in category j ($1 \leq j \leq q$) is $n_j - 1$, and therefore the total number of independent components in the problem of q categories of mixture components is

$$\sum_{j=1}^{q} (n_j - 1) = k - q \quad . \tag{1.3.8}$$

The experimental designs which are convenient to adopt depend greatly on the number k-q of independent components because this number affects the "shape" (number of vertices, edges, faces, etc.) of the experimental region as we show in Section 2.3.

1.4 The general purpose.

In response surface work, it is usually assumed that there is some unknown functional form (1.1.1) which exactly describes the response surface of the system in some region of operation. Response surface methods are employed in an attempt to gain information about the system whether it be locating stationary points or approximating (1.1.1) over some region of interest for the purpose of future response estimation.

The emphasis in this paper will be the estimation of the true functional form (1.1.1) with a polynomial of the

form

$$f(\underset{\sim}{x}) = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} \sum_{j=1}^{k} b_{ij} x_i x_j + \ldots \qquad (1\cdot4\cdot1)$$

in a chosen experimental region of interest when constraints such as (1·2·2), (1·3·1) and (1·3·2) are placed on the components. Also, to facilitate accomplishing the above estimation, it is assumed that the experimenter has some knowledge from which to begin experimentation in terms of a base point or reference level of the mixture components. Then over a sufficiently small region of the factor space, the execution of a program of experiments will be conducted in the "neighborhood" of that point. For example, it may be known that in a blend of the four constituents mentioned in Section 1.3, about 10% of x_1, 40% of x_2, 30% of x_3 and 20% of x_4 will make a suitable product. However, other combinations of the components in the immediate neighborhood of this base point might be of interest for many reasons (cost and availability of constituents, etcetera), and therefore our interest will not be concentrated over the entire factor space as in [16] and [20] but rather in this immediate neighborhood. This region will be called the region of interest.

Hereafter, the above mentioned program of experiments will correspond to an experimental design. We shall also make reference to the construction of an N x k matrix of the variables (components) x_i $(1 \leq i \leq k)$ so that each row

represents a point in k-dimensional space. By virtue of the constraints placed on the components, the dimensionality of this space will be reduced. We denote this matrix by $\underset{\sim}{D}_x$ and call it the design matrix. The N rows of $\underset{\sim}{D}_x$ correspond to the N experiments (points in the factor space of the k variables) to be performed and the k elements in the u^{th} row of $\underset{\sim}{D}_x$ are the levels of the k variables to be used in the u^{th} experiment or trial. The use of such a design will enable us to obtain information about the system by taking observations at the design points. This information will be used to estimate the coefficients (parameters) in the model (1.1.1) by least squares. That is to say, we shall fit a model of the form

$$y_u = \phi(x_{u1}, x_{u2}, \ldots, x_{uk};\ \gamma_1, \gamma_2, \ldots, \gamma_k) + \epsilon_u \quad (1 \le u \le N), \quad (1.4.2)$$

where ϕ is a single-valued function of the known concomitant variables x_i $(1 \le i \le k)$; $\gamma_1, \ldots, \gamma_k$ are unknown parameters and ϵ_u is the random error which follows some frequency distribution. In general the frequency distribution of ϵ_u is not specified. However, the following are usually assumed:

$$\begin{aligned} &\text{(i)} \quad E(\epsilon_u) = 0, \\ &\text{(ii)} \quad \text{Var}\ (\epsilon_u) = \sigma^2, \\ &\text{(iii)} \quad \epsilon_u \text{ and } \epsilon_{u'} \text{ are independently distributed in the probability sense.} \end{aligned} \quad (1.4.3)$$

Since we are primarily interested in the estimation of a chemical response, we shall be concerned with the complete

estimation equation (1.4.1) over the region of interest and not with an investigation of the individual estimated parameters and their variances.

When talking about designs and polynomial models in response surface work, the terms order and degree are synonomous. In this paper however, the term degree will be used throughout, as for example, "in a second-degree design" and "second-degree polynomial".

We shall next discuss a review of literature concerning mixture experimentation. The development of the factor space of the mixture components and an interesting mathematical (topological) formula for determining the number of boundaries of the factor space will also be discussed.

II. LITERATURE REVIEW AND REPRESENTATION OF POLYNOMIALS.

Unlike our definition of the problem where the interest is centered in some specified small region within the factor space of the components, it is sometimes desired to explore the response surface over the entire factor space. Scheffé [20] introduces designs whose points are uniformly spaced over the entire factor space of the components and calls these designs {k,m}-lattice designs. That is to say, in the {k,m}-lattice design, where k is the number of components which comprise the mixtures, and m is the degree of the polynomial used to estimate the response over the simplex factor space, the proportions used for each component have the m+1 equally spaced values

$$x_i = 0,\ 1/m,\ 2/m,\ldots,\ 1 \qquad (1 \le i \le k).$$

In addition, all possible mixtures with these proportions of the components are used.

In this chapter, we shall make reference to {k,m}-lattices and also to the polynomial functions of degree m in the k components which are called {k,m}-polynomials when we refer to the work of Scheffé [20] and Lambrakis [16]. The discussion of the {k,m}-lattices and {k,m}-polynomials however will not be continued after this chapter as we develop new designs and new models for our mixture problem.

2.1 Literature review.

In presenting a method for solving the mixture

problem, Scheffé [20] discusses modifications of his simplex-lattice designs when one or more of the components are restricted by upper bounds. The simplest modification is the replacement of the restricted components with mixtures consisting of combinations of the restricted components and predetermined proportions of the unrestricted components. These mixtures are then used to obtain observations from which estimates of the parameters in his "canonical" polynomials (discussed in Section 2.2) can easily be calculated.

For example, suppose for simplicity only one component is restricted, say $x_1 \leq h$, where the system consists of four components x_1, x_2, x_3, x_4. We shall assume also that we want to fit a second-degree polynomial of the form

$$y(\underset{\sim}{x}) = \sum_{i=1}^{4} \beta_i x_i + \sum_{i<j}^{3} \sum^{4} \beta_{ij} x_i x_j + \epsilon. \qquad (2.1.1)$$

The derivation of the model (2.1.1) is outlined in [20] as well as in Section 2.2. Since the design points in the simplex-lattice arrangement where $x_1 > h$ cannot be used (that is, the design space is now the frustrum of the simplex satisfying $x_1 \leq h$), the points $x_1 > h$ are replaced by combinations consisting of the proportion h of x_1 and the proportions x_j $(2 \leq j \leq 4)$ of the other components where

$$\sum_{j=2}^{4} x_j = 1 - h. \qquad (2.1.2)$$

In this case, if two or more components, other than x_1, are

present, then they are equally represented in Scheffé's scheme, so that, with the notation of (1.3.3), some of the combinations which are used are

$$\begin{aligned} &x_1x_i && (x_1=h,\ x_i=1-h) \\ &x_1x_ix_j && (x_1=h,\ x_i=x_j=\tfrac{1-h}{2}) \\ &x_1x_ix_jx_{j'} && (x_1=\tfrac{h}{2},\ x_i=x_j=x_{j'}=\tfrac{(1-h/2)}{3}) \end{aligned} \qquad (2.1.3)$$

where the subscripts i,j,j' equal 2,3,4 and $i\neq j\neq j'$.

In addition to the design points above, Scheffé also takes the design points where $x_1=0$ and the other components are set at the six combinations of levels ($x_i=1$; $x_i=x_j=1/2$, $j>i>1$) where $x_i=1$ is the i^{th} vertex and $x_i=x_j=1/2$, $j\neq i$ is the midpoint of the edge connecting the i^{th} and j^{th} vertices. The estimates of the parameters β_i, β_{ij} $(j>i>1)$ in (2.1.1) are found in the following manner. If we denote the observed response taken at the mixture $x_i=1$, $x_1=x_j=x_{j'}=0$ $(i\neq j\neq j')$ by y_i and denote the observed response taken at the mixture $x_i=x_j=1/2$, $x_1=x_{j'}=0$ by y_{ij} , then by putting these values of x_1, x_i, x_j, $x_{j'}$, y_i, y_{ij} in (2.1.1) with $E(\underset{\sim}{\varepsilon})=\underset{\sim}{0}$, $E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}')=\sigma^2\underset{\sim}{I}$, we have the least squares estimates

$$\begin{aligned} &\hat{\beta} = y_i && i=2,3,4 \\ &\hat{\beta} = 4y_{ij} - 2y_i - 2y_j && (j>i>1). \end{aligned} \qquad (2.1.4)$$

Of course the estimates of the polynomial coefficients are rather simple linear functions of the observed responses measured at the design points.

Still to be determined are the estimates of β_1 and

β_{1j} (j>1) in (2.1.1). If the observed response taken at the combination $x_1=h$, $x_j=1-h$ (j=2,3,4) is denoted by y'_{1j}, and the observed response taken at the combination $x_1=h/2$, $x_2=x_3=x_4=(1-h/2)/3$ is denoted by y'_1, then the estimates $\hat{\beta}_1$ and $\hat{\beta}_{1j}$ can be found from the following equations,

$$h\hat{\beta}_1 + h(1-h)\hat{\beta}_{1j} = y'_{1j} - (1-h)\hat{\beta}_j \qquad (j=2,3,4) \qquad (2.1.5)$$

$$\frac{h}{2}\hat{\beta}_1 + \frac{h}{2}\frac{(1-h/2)}{3}\sum_{j=2}^{4}\hat{\beta}_{1j} = y'_1 - \frac{(1-h/2)}{3}\sum_{j=2}^{4}\hat{\beta}_j - \frac{(1-h/2)^2}{9}\sum_{i<j}^{3}\sum^{4}\hat{\beta}_{ij}. \qquad (2.1.6)$$

If we solve the equations (2.1.5) and (2.1.6) simultaneously, and this is possible since the $\hat{\beta}_i$, $\hat{\beta}_{ij}$ (i>1, j>i>1) have been previously calculated, we obtain the estimates $\hat{\beta}_1$ and $\hat{\beta}_{1j}$ (j=2,3,4) as we said before.

We notice that Scheffé does not cover the situation where all the components are restricted by upper bounds. That is to say, Scheffé discusses the situation only where one of the components is restricted and merely states that similar modifications for the cases of more than one component remain to be explored. In addition, Scheffé confesses that a shortcoming of the method just described is that the method does not distribute the experimental points very evenly over the part of the factor space satisfying the bounds imposed (see [20], page 355).

Let the components ξ_i ($\xi_i \geq 0$, $\xi_1 + \xi_2 + \ldots + \xi_k = 1$) be

redefined as standardized variables z_i by

$$z_i = \frac{\xi_i - \xi_{0i}}{r_i} \qquad (1 \le i \le k), \tag{2.1.7}$$

where ξ_{0i} ($\xi_{0i} \ge 0$, $\xi_{01}+\xi_{02}+\ldots+\xi_{0k}=1$) denotes the center of the interval of interest for the i^{th} component and $2r_i$ denotes the range of ξ_i chosen by the experimenter. Box and Gardner [5] derive a procedure for solving the problem of placing up to p constraints on the k variables $z_1,\ldots,z_k$ ($p<k$). The procedure involves starting with a design matrix $\underset{\sim}{D}_x$ consisting of a two-level factorial or fractional factorial in k variables. Then with a k x p ($k \ge p$) matrix $\underset{\sim}{W}$ of constraints where the i^{th} column of $\underset{\sim}{W}$ contains the coefficients w_{ui} ($1 \le u \le k$) in the p constraints

$$w_{1i}z_1 + w_{2i}z_2 + \ldots + w_{ki}z_k = 0 \qquad (1 \le i \le p), \tag{2.1.8}$$

the design matrix $\underset{\sim}{D}_z$ of the constrained variables is defined by

$$\underset{\sim}{D}_z = \underset{\sim}{D}_x[\underset{\sim}{I} - \underset{\sim}{W}(\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'], \tag{2.1.9}$$

so that $\underset{\sim}{D}_z\underset{\sim}{W} = \underset{\sim}{0}$. The idempotent matrix $[\underset{\sim}{I} - \underset{\sim}{W}(\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}']$ in (2.1.9) is of the same form as the projection matrix used for obtaining the residual vector in least-squares theory. This same projection principle is used in (2.1.9) in that the elements of the design matrix $\underset{\sim}{D}_z$ are the residuals obtained by "regressing" (projecting) the rows (points in the design space) of a first-degree design matrix $\underset{\sim}{D}_x$ onto the p-dimensional space defined by the constraints.

Now Box and Gardner admit to a difficulty which sometimes occurs. After constructing the matrix of constraints where the experimenter selects the ranges of the ξ_i, the projections sometimes exceed the region of operability (± 1) for the z_i. Consequently the inverse transformation from the z_i to the ξ_i does not always produce operable levels of the ξ_i. In this case, an adjustment on the z_i has to be made. Also, although the method appears to be useful for obtaining a set of design points in working with constrained variables, there is no mention by Box and Gardner of an attempt to optimize the design points for the purpose of estimation, etc. In addition, the procedure is restricted to first-degree models only, a restriction which is clearly worth slackening to second-degree models.

When the variables have upper and lower bounds defined by

$$0 \leq \ell_i \leq x_i \leq \ell_i' \leq 1, \quad (1 \leq i \leq k) \qquad (2.1.10)$$

the factor space is a convex polytope. McLean and Anderson [17] have developed a technique whereby the vertices of the factor space or convex polytope can be located. These points and convex combinations of them can be used as design points with which to estimate a polynomial over the factor space. Once the limits ℓ_i and ℓ_i' are established however, the design is determined and the possibility of obtaining a uniform distribution of design points over the factor space is eliminated. When this non-uniformity of the distribution

of design points exists, that is, when there are clusters of points in some areas of the factor space and only a few points in other areas, it is unlikely that the estimated response function will adequately predict the response over the entire factor space. This is because the variance of the predictor will be affected by the distribution of the design points: poor precision will result in areas of sparse experimentation while there will be good precision in areas with clusters of points. We do not wish to imply that we have constant variance over the entire region of interest, in fact, it would be very reasonable to arrange the variance of $\hat{y}$ to be smallest at the point of maximum interest and increase as we move away from the point. This indeed is approximately attained by means of rotatable designs.

Probably the work that most closely parallels our form of the mixture problem is the work of Lambrakis [16] in which the components satisfying the constraints

$$x_i \geq 0 , \quad x_1 + x_2 + \ldots + x_n = 1 ,$$

are defined as "major" components. These components are themselves mixtures of one or more "minor" components. The "major" components or as we shall refer to them, M-components, are defined as having the proportions

$$c_i \geq 0 \qquad (1 \leq i \leq n) \tag{2.1.11}$$

in a mixture so that with n M-components,

$$c_1 + c_2 + \ldots + c_n = 1. \tag{2.1.12}$$

Let n_i be the number of components in the i^{th} M-component. The proportion of the j^{th} component $(1 \le j \le n_i)$ in the i^{th} M-component is defined by

$$x_{ij} \ge 0 \qquad (1 \le i \le n,\ 1 \le j \le n_i),$$

and

$$\sum_{j=1}^{n_i} x_{ij} = 1 \qquad \text{for each } i = 1,2,\ldots,n. \tag{2.1.13}$$

If we combine (2.1.11), (2.1.12) and (2.1.13), we see that the proportion of the j^{th} component in the i^{th} M-component present in the mixture is

$$X_{ij} = c_i x_{ij} \qquad (1 \le i \le n,\ 1 \le j \le n_i), \tag{2.1.14}$$

and these proportions satisfy the condition that

$$\sum_{i=1}^{n} \sum_{j=1}^{n_i} X_{ij} = 1. \tag{2.1.15}$$

Lambrakis discusses the problem involving just two M-components and suggests the extension to n M-components. The c_i $(1 \le i \le 2)$ are assumed to be fixed with non-zero values, although Lambrakis does not specify the values of c_i as we did by setting $c_i = \frac{1}{q}$ (see (1.3.1) and (1.3.2)). The proportions x_{ij} in each M-component take the values

$$x_{ij} = 0, \frac{1}{m_i}, \frac{2}{m_i}, \ldots, 1 \qquad (1 \le i \le 2) \tag{2.1.16}$$

where m_i is the degree of the polynomial used in estimating the response over the i^{th} $(1 \le i \le 2)$ simplex lattice, that is, the simplex space corresponding to the i^{th} M-component.

For example, if $m_i=2$, then

$$x_{ij} = 0, \frac{1}{2}, 1 \qquad (1\leq j\leq n_i),$$

if $m_i=3$,

$$x_{ij} = 0, \frac{1}{3}, \frac{2}{3}, 1 \qquad (1\leq j\leq n_i).$$

The m_i $(1\leq i\leq n)$ do not have to be equal, that is to say, one may consider fitting polynomials of different degrees over the individual M-component lattices (by "fitting a polynomial", we mean the construction of the lattice arrangement of design points which is appropriate for fitting a polynomial of some specified degree), and in fact, Lambrakis considers the case where $m_1=3$ and $m_2=2$.

The arrangement of design points suggested by Lambrakis is the result of combining the n $\{n_i,m_i\}$-lattices. The resulting lattice arrangement is called a multiple lattice and the notation used is $\{n_1,\ldots,n_n;m_1,\ldots,m_n\}$-multiple lattice. For example, when n=2, $m_1=3$ and $m_2=2$, the arrangement of design points is called a $\{n_1,n_2;3,2\}$-double lattice. The mixture combinations (design points) in the double lattice are obtained by combining the design points in each of the individual M-component lattices. That is, if $n_1=2$, $n_2=2$, $m_1=2$, $m_2=2$ and X_1, X_2 are the components in M-component 1 and X_3, X_4 are the components in M-component 2, then the mixture combinations in the $\{2,2;2,2\}$-double lattice are

$$X_iX_j \quad (i=1,2;j=3,4), \qquad X_iX_3X_4 \quad (i=1,2),$$
$$X_1X_2X_j \quad (j=3,4), \qquad X_1X_2X_3X_4 \qquad (2.1.17)$$

where in (2.1.17), $X_i=c_1$ $(x_i=0,\frac{1}{2},1)$ for i=1,2; $X_j=c_2$ $(x_j=0,\frac{1}{2},1)$ for j=3,4 and in the combination $X_iX_3X_4$, $x_i=1$, $x_3=\frac{1}{2}$, $x_4=\frac{1}{2}$. If $c_1=c_2=\frac{1}{2}$, the {2,2;2,2}-double lattice arrangement is shown in Figure 2 but with a center point added.

As we stated before, Lambrakis assumes the interest is over the entire factor space of the components. Consequently, the lattice arrangement of design points will differ from the arrangement of design points that we suggest in the following chapters where our interest is isolated to some smaller region within the factor space. In addition, the polynomial models used for estimating the response over the two (ours and his) regions of interest will be of different forms. We shall discuss these representations of the polynomial models in Section 2.2.

For the original (Scheffé) mixture problem, the use of response surface designs to estimate continuous polynomial models over a specified region of interest was first introduced by Thompson and Myers [22]. Thompson and Myers define as the region of interest an ellipsoidal region which is contained entirely in the simplex factor space and is centered about a point of maximum interest. The definition of the ellipsoidal region of interest enables the experimenter to define symmetric intervals of interest, not necessarily all equal, for each of the mixture components. The point of maximum interest provides a convenient starting point for the

experimentation as well as a base point from which to construct designs.

Thompson and Myers show, with the use of derived power vectors and Schläflian matrices, how polynomial models of any degree can be used to estimate the response over the region of interest. A transformation from the set of k mixture components to a set of k-1 linearly independent variables is made. The transformation to independent variables enables the use of standard methods of design construction as well as facilitates the use of the criterion, the average mean square error of the predictor $\hat{y}$, for determining optimal design configurations. A similar transformation to independent variables and the definition of an ellipsoidal region of interest will be discussed in Chapter 3 since an approach, similar to the approach of Thompson and Myers, is undertaken in this paper.

2.2 Representations of polynomials.

In Section 1.4, it was mentioned that for prediction purposes, a polynomial equation in the k components $x_1,\ldots,x_k$ which can be written as

$$f(\underline{x}) = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{1\le i_1\le i_2}^{k}\sum^{k} b_{i_1 i_2} x_{i_1} x_{i_2}$$

$$+ \ldots + \sum_{1\le i_1\le\ldots\le i_d}^{k}\ldots\sum^{k} b_{i_1\ldots i_d} x_{i_1}\ldots x_{i_d} \qquad (2.2.1)$$

can be used. The number of terms in equation (2.2.1) is $\binom{k+d}{d}$. Since the components x_i $(1 \leq i \leq k)$ in equation (2.2.1) represent proportions of the k constituents in a mixture and thus must sum to unity, we may make the substitution

$$x_k = 1 - \sum_{i=1}^{k-1} x_i \qquad (2.2.2)$$

in (2.2.1) and this does not affect the degree of the polynomial. The effect of substituting (2.2.2) in (2.2.1) is that now $f(\underline{x})$ becomes a polynomial of degree d in the k-1 components $x_1, x_2, \ldots, x_{k-1}$ with $\binom{k-1+d}{d}$ terms. Thus the model is simpler than before; but the effect of the k^{th} component is not expressed in the polynomial and hence is obscured by the formula.

In [20], Scheffé derives the "canonical" polynomials by applying the constraint (1.2.2) to the terms in the general polynomial (2.2.1). The canonical polynomials possess the property that they are easy to use for prediction purposes over the simplex factor space and also contain the same number of terms as there are points in the associated designs. For example, the quadratic model (see (2.1.1)) for a system containing four components requires ten terms, and the second-degree lattice design contains ten mixture points. However, since the canonical polynomials contain the same number of terms as there are points in the designs, there is no way to determine the lack of fit of the models without

augmenting the designs. When the designs are augmented, they lose their simplicity.

Since the substitution principle used for deriving the canonical polynomials can be applied to our problem, we now demonstrate the construction of the canonical polynomials, after which we shall derive the polynomials for our problem. Suppose d=1, then equation (2.2.1) is

$$f(\underset{\sim}{x}) = b_0 + \sum_{i=1}^{k} b_i x_i \ .$$

If we multiply the constant term b_0 by $x_1+x_2+\ldots+x_k=1$, then

$$f(\underset{\sim}{x}) = b_0\left(\sum_{i=1}^{k} x_i\right) + \sum_{i=1}^{k} b_i x_i$$

$$= \sum_{i=1}^{k} c_i x_i$$

where $c_i = b_0 + b_i$. Note, when constructing polynomials, the c_i correspond to parameter estimates and are not used in the same sense as in Section 2.1. For a second-degree model,

$$f(\underset{\sim}{x}) = b_0 + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i^{\,2} + \sum\sum_{i<j}^{k} b_{ij} x_i x_j \ .$$

If we use the identity

$$x_i^{\,2} = x_i\left(1 - \sum_{\substack{j=1 \\ j\neq i}}^{k} x_j\right) \ ,$$

then

$$f(\underset{\sim}{x}) = b_0 \sum_{i=1}^{k} x_i + \sum_{i=1}^{k} b_i x_i + \sum_{i=1}^{k} b_{ii} x_i \Big(1 - \sum_{\substack{j=1 \\ j\neq i}}^{k} x_j\Big) + \sum_{i<j}^{k}\sum b_{ij} x_i x_j$$

$$= \sum_{i=1}^{k} (b_0 + b_i + b_{ii}) x_i - \sum_{i=1}^{k} b_{ii} x_i \sum_{\substack{j=1 \\ j\neq i}}^{k} x_j + \sum_{i<j}^{k}\sum b_{ij} x_i x_j$$

$$= \sum_{i=1}^{k} c_i x_i + \sum_{i<j}^{k}\sum c_{ij} x_i x_j \qquad (2.2.4)$$

where $c_i = b_0+b_i+b_{ii}$ and $c_{ij} = b_{ij}-b_{ii}-b_{jj}$. Furthermore, the equation (2.2.4) can also be written in the homogeneous form

$$f(\underset{\sim}{x}) = \sum_{i=1}^{k} d_{ii} x_i^{\,2} + \sum_{i<j}^{k}\sum d_{ij} x_i x_j$$

$$= \sum_{i\leq j}^{k}\sum_{k} d_{ij} x_i x_j \qquad (2.2.5)$$

by letting

$$\sum_{i=1}^{k} c_i x_i = \sum_{i=1}^{k} c_i x_i \Big(\sum_{j=1}^{k} x_j\Big)$$

and simplifying. We see then that if we apply the constraint $x_1 + x_2 + \ldots + x_k = 1$ to the terms in the general polynomial model, the polynomial model can be written in many forms. We shall discuss the result of this non-uniqueness property of the model later when matters of design criteria arise.

A polynomial of the d^{th} degree in k components where

the components belong to q categories can be reduced to the form

$$f(\underset{\sim}{x}) = c_0 + \sum_{\substack{i=1 \\ i\neq n_j}} c_i x_i + \sum \sum_{\substack{i\leq i_1\leq i_2 \\ i_1,i_2\neq i_{n_j}}} c_{i_1 i_2} x_{i_1} x_{i_2} + \ldots$$

$$+ \sum_{\substack{1\leq i_1\leq\ldots\leq i_d \\ i_1,\ldots,i_d\neq i_{n_j}}} \ldots \sum c_{i_1\ldots i_d} x_{i_1}\ldots x_{i_d}\,, \qquad (2.2.6)$$

by making the substitution

$$x_{n_j} = \frac{1}{q} - \sum_{i=n_{j-1}+1}^{n_j-1} x_i \qquad (1\leq j\leq q) \qquad (2.2.7)$$

in (2.2.1) and remembering identities such as

$$x_i^2 = x_i\left(\frac{1}{q} - \sum_j x_j\right) \qquad (n_{j-1}+1\leq j\leq n_j;\ j\neq i).$$

As in (1.3.2), n_j in (2.2.7) is the number of components in the j^{th} category. As we previously mentioned in this section however, the substitution of (2.2.7) in (2.2.1) resulting in (2.2.6) does not give the experimenter any simple measure of the effects of the n_j^{th} $(1\leq j\leq q)$ components and therefore a different approach will be used.

An approach identical to the one used by Scheffé in deriving the canonical polynomials can be taken and we

shall show the model as well as the geometrical representation of the factor space for the case where q=2, k=4.

If we recall from (2.2.4), the Scheffé quadratic model in four components is written as

$$\eta(x) = \sum_{i=1}^{4} \beta_i x_i + \sum_{i<j}^{3} \sum^{4} \beta_{ij} x_i x_j \ .$$

Now if we apply the constraints

$$x_1 + x_2 = \frac{1}{2} \ , \qquad x_3 + x_4 = \frac{1}{2} \tag{2.2.8}$$

and at the same time repeatedly multiply all the terms in the model by $x_1+x_2+\ldots+x_k=1$, then the above model can be rewritten as

$$\begin{aligned} \eta = \sum_{i=1}^{2} \sum_{j=3}^{4} & \left[2(1-\lambda)(\beta_i+\beta_j) + \beta_{ij} \right] x_i x_j \\ & + 2(1-\lambda)\beta_{12}(x_1x_2x_3 + x_1x_2x_4) \\ & + 2(1-\lambda)\beta_{34}(x_1x_3x_4 + x_2x_3x_4) \\ & + \lambda\left[\sum_{i=1}^{4} \beta_i x_i + \beta_{12}x_1x_2 + \beta_{34}x_3x_4 \right], \end{aligned} \tag{2.2.9}$$

where $\lambda = 1/2^n$. In fact this is true for any real number λ as we see by letting n=0, ∞. That is to say, if n=0 then λ=1 and (2.2.9) is

$$\eta = \sum_{i=1}^{2} \sum_{j=3}^{4} \beta_{ij} x_i x_j + \beta_{12}x_1x_2 + \beta_{34}x_3x_4 + \sum_{i=1}^{4} \beta_i x_i .$$

If $n=\infty$, $\lambda=0$, then

$$\eta = \sum_{i=1}^{2} \sum_{j=3}^{4} \left[2(\beta_i + \beta_j) + \beta_{ij} \right] x_i x_j + 2\beta_{12}(x_1x_2x_3 + x_1x_2x_4)$$

$$+ 2\beta_{34}(x_1x_3x_4 + x_2x_3x_4).$$

If we substitute the constraints (2.2.8) in the above model, the result is

$$\eta = \sum_{i=1}^{4} \beta_i x_i + \sum\sum_{i<j}^{4} \beta_{ij} x_i x_j .$$

Since (2.2.9) is true for two values of λ, it must be true for all values of λ.

The model (2.2.9) can be written in a simpler form by putting $\lambda=0$ and replacing the coefficients β_{ij}, β_i, β_j with γ_{ij}, γ_{ijk}. That is, if

$$\gamma_{ij} = 2(\beta_i + \beta_j) + \beta_{ij}$$

$$\gamma_{123} = \gamma_{124} = 2\beta_{12}, \qquad \gamma_{134} = \gamma_{234} = 2\beta_{34},$$

then equation (2.2.9) is

$$\eta = \sum_{i=1}^{2} \sum_{j=3}^{4} \gamma_{ij} x_i x_j + \sum_{i=1}^{2} \gamma_{i34} x_i x_3 x_4$$

$$+ \sum_{j=3}^{4} \gamma_{12j} x_1 x_2 x_j \qquad (2.2.10)$$

and the estimates $\hat{\gamma}_{ij}$, $\hat{\gamma}_{ij\ell}$ of the parameters in (2.2.10) will be linear functions of the observations. Also, we see that the individual terms in the model (2.2.10) are combinations of the components which produce valid mixtures, and

thus the model appears to represent the problem better than does the form (2.2.6).

We have been talking about using a quadratic model in the four mixture components. Therefore, each component must be set at at least three levels. Put $x_i=0, \frac{1}{4}, \frac{1}{2}$ (i=1,2,3,4). The design configuration (we omit the center point) is shown in Figure 2 on the next page. Note that the number of terms (namely eight) in (2.2.11) corresponds exactly to the number of design points in Figure 2. The result is, like the Scheffé canonical polynomials, this particular form of the model does not permit the experimenter to measure the lack of fit of the model and therefore the degree of accuracy of the predictions using this model cannot be determined. By adding a center point, we can obtain some measure of the lack of fit of the model, but the total number of design points is now nine.

Lambrakis uses an interesting approach for constructing a regression function which is to be used for representing the response over the multiple-lattice design space. He conjectures that one can combine the terms in each of the canonical polynomials where each polynomial would be used separately to estimate the response over the simplex space corresponding to each of the n M-components. That is to say, if n=2, the appropriate regression function is formed by multiplying together the two polynomials. For example, let n=2, $n_1=n_2=2$ and $m_1=m_2=2$. If x_1, x_2 comprise M-component

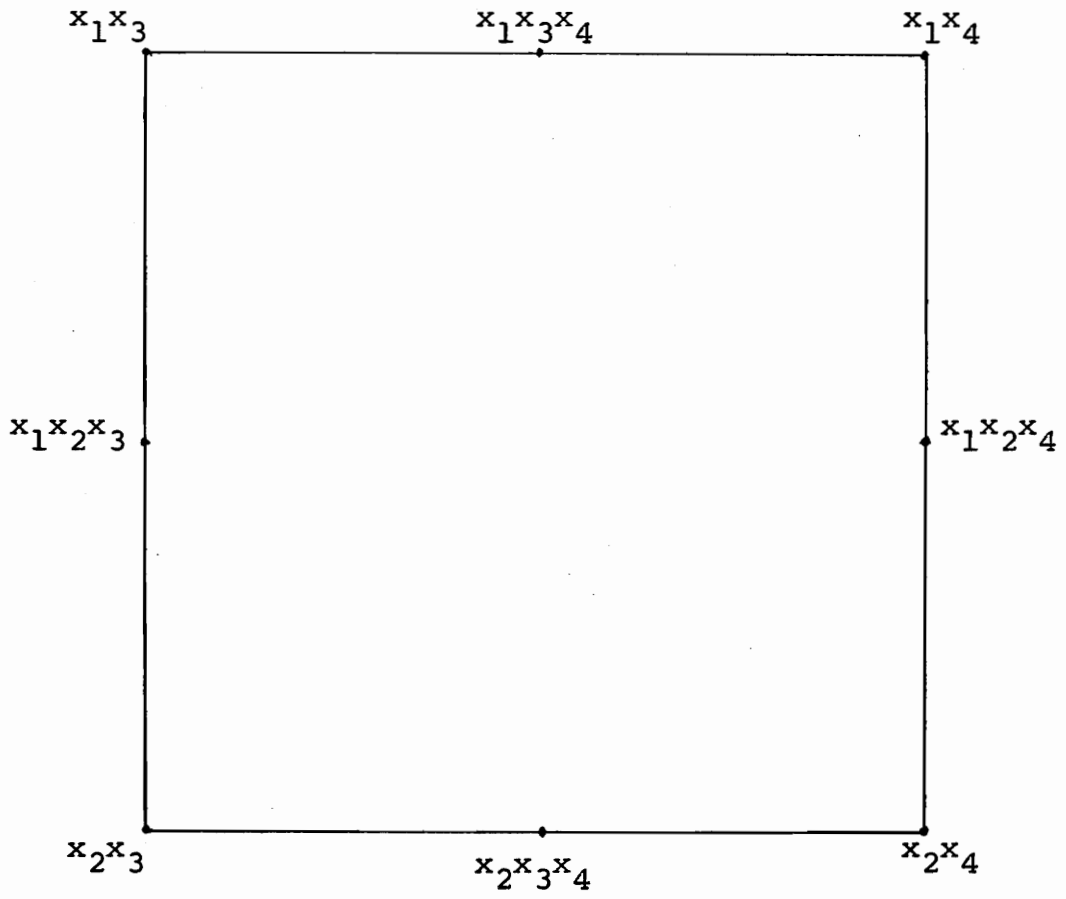

Figure 2.

Mixture Combinations for the Quadratic Model in the Four Components.

1 and x_3 and x_4 comprise M-component 2, then from (2.2.4), we can let

$$\eta_1 = \sum_{i=1}^{2} d_i x_i + d_{12} x_1 x_2$$

and

$$\eta_2 = \sum_{j=3}^{4} d_j x_j + d_{34} x_3 x_4$$

represent the second-degree models for the M-components 1 and 2 respectively. Then the regression function which would be used to represent the response over the {2,2;2,2}-double lattice is

$$\eta = \sum_{i=1}^{2} \sum_{j=3}^{4} \gamma_{ij} x_i x_j + \sum_{i=1}^{2} \gamma_{i34} x_i x_3 x_4 + \sum_{j=3}^{4} \gamma_{12j} x_1 x_2 x_j + \gamma_{1234} x_1 x_2 x_3 x_4, \qquad (2.2.11)$$

and the estimates $\hat{\gamma}_{ij}$, $\hat{\gamma}_{i34}$, $\hat{\gamma}_{12j}$, $\hat{\gamma}_{1234}$ are linear functions of the observations taken at the points in the double lattice. The model (2.2.11) is identical to (2.2.10) except for the additional term $\gamma_{1234} x_1 x_2 x_3 x_4$ which corresponds to the center point of the {2,2;2,2}-double lattice design. The number of terms in (2.2.11) is nine which corresponds to the number of design points in the double lattice. This is one disadvantage of the use of lattice designs as well as of the use of these models. That is, we show later that designs containing fewer points and models containing fewer terms can be used for exploring the response surface over a

two-dimensional factor space.

With lattice designs, the choice of the particular arrangement of design points as in Figure 2 is made under the assumption that the response is fairly smooth over the entire factor space. Consequently, ridges or depressions in the surface which occur within the bounds of the factor space may not be correctly interpreted. Also, with the experimental arrangement in Figure 2 as well as Lambrakis' double-lattice designs, although the variance of $\hat{y}$ for a linear model is minimized because of the spread of the design points, the variance of the predictor $\hat{y}$ for a second-degree model is not necessarily constant at a given distance from the center of the design. Thus the experimenter takes a risk with his predictions at different locations in the factor space which are equidistant from the center of the design. This concept will be discussed at further length in Chapter 5 when we suggest the use of rotatable designs.

Another objection to the use of lattice arrangements of design points is that some mixture combinations do not contain all the components. (This objection to the use of purely lattice arrangements was recognized by Draper and Lawrence [10].) For example, in Figure 2, the four design points (mixture combinations) at the vertices of the square consist of only two components and the exploration is over a system involving four components. This loss of information will be remedied with our designs in that every design

point of our designs will contain all of the components unless otherwise chosen by the experimenter.

2.3 Shape of the convex polytope.

Probably the major impact of the additional constraints (1.3.1) and (1.3.2) placed on the mixture components is that the (k-1)-dimensional simplex of Scheffé's mixture problem can no longer be used to represent the factor space of the k components. Instead, with q categories of k mixture components, the factor space is a (k-q)-dimensional convex polytope. Lambrakis describes the factor space as a multiple-simplex. The description of the polytope might best be explained by the following example where q=2.

The general shape of the convex polytope, that is, the number of extremities (vertices, edges, faces, etc.) will of course depend on the number of components in each category. Let us designate the number of components in category 1 by n where the total number of components is k and $k \geq n+2$. When no constraints are placed on the components, that is, with k independent components, the factor space is usually called k-space. With the constraint (1.2.2), the dimensionality of the factor space is reduced by 1 and as we pointed out before, by the nature of the constraint $(x_1+x_2+\ldots+x_k=1)$, the factor space can be geometrically represented by a (k-1)-dimensional regular simplex. With the

additional constraint $x_1+x_2+\ldots+x_n=\frac{1}{2}$ of our problem, the dimensionality is reduced still further to k-2. For each additional constraint placed on the components, the dimensionality of the figure is reduced by 1. For a general discussion of the geometry of sections and frusta of simplices, see Appendix B.

Let x_i (i=1,2,...,n) denote the components in category 1 and x_j (j=n+1,...,k) the components in category 2. If n=2 and k=4, then by the constraints

$$x_1 + x_2 = \frac{1}{2}\,, \qquad x_3 + x_4 = \frac{1}{2}\,, \tag{2.3.1}$$

we know that the factor space in two-dimensions is square by virtue of symmetry arguments. With an additional constraint $x_i=0$ (i=1,2), the dimensionality of the figure is reduced to 1 and we can infer that

$$x_i = 0,\ x_{i'} = \frac{1}{2}\,, \quad x_3 + x_4 = \frac{1}{2} \quad (i'=1,2;i'\neq i), \tag{2.3.2}$$

is a line or edge of the square defined by (2.3.1). With the constraint $x_j=0$ (j=3,4) added to (2.3.2), the combination

$$x_i = 0 \quad x_{i'} = \frac{1}{2} \quad x_j = 0 \quad x_{j'} = \frac{1}{2} \quad (i\neq i';j\neq j'), \tag{2.3.3}$$

denotes a vertex of the square. Note that in applying the constraints in (2.3.3), we could not have set $x_i=0$, $x_{i'}=0$ (i=1,2;i'=1,2;i≠i'), $x_3+x_4=1$ nor $x_1+x_2=1$, $x_j=0$, $x_{j'}=0$ (j=3,4; j'=3,4;j≠j') and still have a valid mixture, that is, satisfy the original constraints $x_1+x_2=x_3+x_4=1/2$.

In order to establish a method for determining the number and also give some description of the extremities of

the convex polytope, we shall work through a larger example where n=2, k=6. If we let μ be the number of additional constraints over and above the original constraints (where a constraint refers to the setting of at least one $x_i=0$) placed on the components, then for

$\mu = 1$

$x_1 = 0 \quad x_2 = \frac{1}{2} \quad x_3+x_4+x_5+x_6 = \frac{1}{2}$

$x_1 = \frac{1}{2} \quad x_2 = 0 \quad x_3+x_4+x_5+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_3 = 0 \quad x_4+x_5+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_4 = 0 \quad x_3+x_5+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_5 = 0 \quad x_3+x_4+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_6 = 0 \quad x_3+x_4+x_5 = \frac{1}{2}$

six hyperplanes (dimensionality 3)

$\mu = 2$

$x_1 = 0 \quad x_2 = \frac{1}{2} \quad x_3 = 0 \quad x_4+x_5+x_6 = \frac{1}{2}$

$x_1 = 0 \quad x_2 = \frac{1}{2} \quad x_4 = 0 \quad x_3+x_5+x_6 = \frac{1}{2}$

$x_1 = 0 \quad x_2 = \frac{1}{2} \quad x_5 = 0 \quad x_3+x_4+x_6 = \frac{1}{2}$

$x_1 = 0 \quad x_2 = \frac{1}{2} \quad x_6 = 0 \quad x_3+x_4+x_5 = \frac{1}{2}$

$x_1 = \frac{1}{2} \quad x_2 = 0$ same four as above

$x_1+x_2 = \frac{1}{2} \quad x_3 = x_4 = 0 \quad x_5+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_3 = x_5 = 0 \quad x_4+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_3 = x_6 = 0 \quad x_4+x_5 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_4 = x_5 = 0 \quad x_3+x_6 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_4 = x_6 = 0 \quad x_3+x_5 = \frac{1}{2}$

$x_1+x_2 = \frac{1}{2} \quad x_5 = x_6 = 0 \quad x_3+x_4 = \frac{1}{2}$

fourteen faces (dimensionality 2)

$\mu = 3$	$x_1 = 0$	$x_2 = \frac{1}{2}$	same six as above	
	$x_1 = \frac{1}{2}$	$x_2 = 0$	same six as above	
	$x_1+x_2 = \frac{1}{2}$	$x_3 = x_4 = x_5 = 0$	$x_6 = \frac{1}{2}$	sixteen edges
	$x_1+x_2 = \frac{1}{2}$	$x_3 = x_4 = x_6 = 0$	$x_5 = \frac{1}{2}$	(dimensionality 1)
	$x_1+x_2 = \frac{1}{2}$	$x_3 = x_5 = x_6 = 0$	$x_4 = \frac{1}{2}$	
	$x_1+x_2 = \frac{1}{2}$	$x_4 = x_5 = x_6 = 0$	$x_3 = \frac{1}{2}$	
$\mu = 4$	$x_1 = 0$	$x_2 = \frac{1}{2}$	same four as above	eight vertices.
	$x_1 = \frac{1}{2}$	$x_2 = 0$	same four as above	(dimensionality 0)

The above enumeration of the boundaries can be checked by a generalization of Euler's formula, namely that for a simply connected polytope of any number of dimensions, $\tilde{N}_0-\tilde{N}_1+\tilde{N}_2-\ldots=0$ or 2, depending on whether the number of dimensions of the polytope is even or odd, where $\tilde{N}_r$ is the number of r-flats bounding the polytope. This useful check does not seem to be in the literature of response surfaces for mixtures, although it is well known to topologists, see for example, Coxeter, H.S.M. (1963), Regular Polytopes, Second edition, The MacMillan Company, New York,p.9-11,58-59; Manning, M.P. (1956), Geometry of Four Dimensions, Dover Publications, Inc., New York,p.300-306.

We are now ready to generalize the method for determining the number of extremities. Let μ be the number of constraints ($1\leq\mu\leq k-2$) placed on $x_1+\ldots+x_n=x_{n+1}+\ldots+x_k=\frac{1}{2}$ and also let $m=k-n$ where $m\geq n$. If M and Q are integers and given $\binom{M}{Q}$ is zero when either $Q>M$ or Q is negative, the

number $\tilde{N}$ of extremities of dimensionality $k-2-\mu$ of the convex polytope for the cases: $n=2$ $k=4,5,6,7$; $n=3$ $k=6,7,8$; $n=4$ $k=8,9$ is given by

$$\tilde{N} = \binom{k}{\mu} - \left[\binom{m}{\mu-n} + \binom{n}{\mu-m}\right] . \qquad (2.3.4)$$

The validity of (2.3 4) can be verified by considering when $\mu=n<m$,

$$\tilde{N} = \binom{n}{n-1}\binom{m}{1} + \binom{n}{n-2}\binom{m}{2} + \ldots + \binom{n}{0}\binom{m}{n} .$$

When $n=\mu=m$,

$$\tilde{N} = \binom{n}{n-1}\binom{m}{1} + \ldots + \binom{n}{1}\binom{m}{m-1} .$$

When $\mu<n<m$,

$$\tilde{N} = \binom{n}{\mu}\binom{m}{0} + \binom{n}{\mu-1}\binom{m}{1} + \ldots + \binom{n}{0}\binom{m}{\mu},$$

etc.

The generalization of Euler's formula is

$$\tilde{N}_0 - \tilde{N}_1 + \tilde{N}_2 - \ldots + (-1)^{r-1}\tilde{N}_{r-1} = 1 - (-1)^r \qquad (2.3.5)$$

where $\tilde{N}_\ell$ is the number of ℓ-dimensional boundaries and r ($\ell \leq r$) is the dimensionality of the polytope. If we sum μ in (2.3.4) over the range $1 \leq \mu \leq k-2$ where $k-2=r$ in (2.3.5), we have

$$\sum_{\mu=k-2}^{1} (-1)^{k-2-\mu} \left\{\binom{k}{\mu} - \left[\binom{m}{\mu-n} + \binom{n}{\mu-m}\right]\right\} =$$

$$-\left[(-1)^{k-2-\mu} \left\{\binom{k}{\mu} - \left[\binom{m}{\mu-n} + \binom{n}{\mu-m}\right]\right\}\right]_{\mu=0,k-1,k.} \qquad (2.3.6)$$

Putting these values of μ in the right-hand side of (2.3.6), we have

$$-\left[(-1)^{k-2} - k + k - 1\right] = 1 - (-1)^{k-2}$$

which agrees with the right-hand side of (2.3.5), that is, the generalization of Euler's formula checks (2.3.4) also.

We next discuss the set-up of the problem by defining a region of interest in the system of the mixture components. We shall make a transformation from the set of mixture components to a set of independent variables so as to be able to conform to standard (response surface) methods of design construction, etc. The transformation to the independent variables also enables us to be able to describe the response system with fewer than k variables.

III. TRANSFORMATION TO INDEPENDENT VARIABLES.

When using polynomial models in the mixture problem, it was shown in Section 2.2 that polynomials used for representing the same response can take on many forms. Although the estimate of the response will be unique for different representations of the polynomial, criteria for optimizing designs cannot be conveniently based on functions of the parameters in the model. This is because the criteria assumed take different forms for different representations of the polynomial and to avoid having to specify the various forms of the criteria for the different representations of the model, the focus of attention will be placed on the predictor $\hat{y}$.

A criterion such as the average mean square error of $\hat{y}$ proposed by Box and Draper [3] can be easily used for constructing optimal designs when the variables are independent. This will be clear when we discuss in Chapter 4 the division of the average mean square error of $\hat{y}$ into two parts; the average variance of $\hat{y}$ and the average squared bias of $\hat{y}$. An additional reason for preferring to work with independent variables is the ease in applying known response surface techniques to the problem. This is because most of the theory of response surface methods (design construction as well as the seeking of optimal operating conditions) has been developed under the assumption that the variables are independent.

In the first chapter (Section 1.3), we mentioned that with q categories of k mixture components, the number of independent components is k-q. Thus the rank of the N x k ($N \geq k$) matrix $\underset{\sim}{X}$ (the form of which will be shown shortly) is k-q in the general linear model

$$\underset{\sim}{y} = \underset{\approx}{X}\ \underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} \qquad (3\cdot0\cdot1)$$

where $\underset{\sim}{y}$ is an N-dimensional random vector of observations, $\underset{\sim}{\gamma}$ is a k x 1 vector of unknown parameters and $\underset{\sim}{\varepsilon}$ is a random error vector ($E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$, $E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \sigma^2 \underset{\approx}{I}_N$). The parameters in equation (3·0·1) must be estimated in order to predict the response for different values of $x_1, x_2, \ldots, x_k$. However, a difficulty exists in that the least-squares estimates $\hat{\underset{\sim}{\gamma}}$ of the $\underset{\sim}{\gamma}$ are not unique since the rank of the matrix $\underset{\sim}{X}$ in (3·0·1) is k-q. Consequently, not only will different sets of estimates $\hat{\gamma}_j$ ($1 \leq j \leq k$) provide the same value for the predictor $\hat{y}$ of y but the usual interpretation of the estimates of the coefficients becomes obscured.

To alleviate this problem of non-uniqueness of the estimates as well as to be able to apply known response surface methods, we shall transform the dependent mixture components to a set of independent variables. However, before we make this transformation, let us consider the nature of the components x_j ($1 \leq j \leq k$) in some typical situations. In one application of a design, x_j may refer to a particular blending ingredient in a gasoline and in another application, to a different type of chemical reagent. It may be desirable

therefore to redefine our general mixture problem in terms of intermediate variables v_j $(1 \leq j \leq k)$ which will be without physical dimensions. We call these variables intermediate variables because they comprise the intermediate step between the mixture components and the independent design variables which we introduce in Section 3.2. We next consider this transformation of the mixture components.

3.1 Transforming the mixture components: Defining the region of interest.

If we turn our attention to the actual setting-up of the experiments, we stated in Section 1.4 that one assumption we are making in this paper is that the experimenter presumably has some knowledge of a location for a base point in the factor space of the mixture components. This base point which will be located in the center of the region of interest (and therefore will specify the center of our design configurations) could be some known desirable mixture or just some convenient point upon which to begin experimentation. In the event that this is an initial attempt in experimentation, a convenient point for the j^{th} $(1 \leq j \leq k)$ component might be $x_{0j} = \frac{1}{qn_j}$, where x_j is from a category which contains n_j components. This will give equal importance with regard to location within the j^{th} category in the factor space to each of the n_j components in the j^{th} category.

With the location of the base point specified, let us now make the transformation involving the mixture components x_j $(1 \leq j \leq k)$ and the base point, which we denote in vector notation by $\underset{\sim}{x}_0' = (x_{01}, x_{02}, \ldots, x_{0k})$. That is to say, the distance $x_j - x_{0j}$ scaled by the positive quantity h_j $(0 < h_j \leq x_{0j})$ defines the symmetric interval of interest for the j^{th} $(1 \leq j \leq k)$ component. Then it will be assumed that the ellipsoidal region

$$\sum_{j=1}^{k} \left(\frac{x_j - x_{0j}}{h_j}\right)^2 \leq 1 \qquad (3.1.1)$$

defines the region of interest for the particular application. Each x_{0j} and h_j $(1 \leq j \leq k)$ in (3.1.1) is chosen by the experimenter so as to give appropriate location and spread $(0 \leq x_{0j} - h_j,\ x_{0j} + h_j \leq \frac{1}{q}$ for all $j = 1,2,\ldots,k)$ to each of the k components. Note that the principal axes of the ellipsoidal region will be parallel to the axes of the mixture components. By defining the region in this way, the transformation which would otherwise be necessary to arrive at the form (3.1.1), i.e., the transformation needed to eliminate the crossproduct terms usually present in the equation of a quadric, is conveniently omitted. Also, as pointed out in [5] and [22], this definition of a reduced region of interest is common in blending problems where the proportions which are needed may only be roughly known.

Let us define the intermediate variables by

$$v_j = \frac{x_j - x_{0j}}{h_j} \qquad (1 \leq j \leq k). \qquad (3\cdot1\cdot2)$$

The ellipsoidal region (3·1·1) is now transformed to the interior and surface of a unit sphere in the metric of the intermediate variables v_j. Also, the unit spherical region will be centered at $v_j = 0$ $(1 \leq j \leq k)$. In Section 4.3, another region defined as the largest sphere centered at $\underset{\sim}{v}=\underset{\sim}{0}$ that will fit inside the convex polytope is derived and this extended region enables us to extend our interest outside the unit spherical region.

In matrix notation, (3·1·2) can be expressed as

$$\underset{\sim}{v} = \underset{\sim}{H}^{-1} (\underset{\sim}{x} - \underset{\sim}{x}_0) \qquad (3\cdot1\cdot3)$$

where $\underset{\sim}{H}$ = diagonal $(h_1, h_2, \ldots, h_k)$. Let the transpose $\underset{\sim}{v}'$ of $\underset{\sim}{v}$ in (3·1·3) define a row, say that u^{th} row of an N x k coefficient matrix $\underset{\sim}{V}$, that is,

$$\underset{\sim}{v}_u' = (\underset{\sim}{x}_u - \underset{\sim}{x}_0)' \, H^{-1} \qquad (1 \leq u \leq N). \qquad (3\cdot1\cdot3A)$$

If the N x k matrix $\underset{\sim}{X}$ in (3·0·1) consists of the elements $x_{uj} - x_{0j}$ $(1 \leq u \leq N,\ 1 \leq j \leq k)$, that is,

$$\underset{\sim}{X} = \begin{bmatrix} x_{11} - x_{01} & x_{12} - x_{02} & \cdots & x_{1k} - x_{0k} \\ x_{21} - x_{01} & x_{22} - x_{02} & \cdots & x_{2k} - x_{0k} \\ & & & \\ x_{N1} - x_{01} & x_{N2} - x_{02} & \cdots & x_{Nk} - x_{0k} \end{bmatrix}, \qquad (3\cdot1\cdot3B)$$

and $\underset{\sim}{v}_u'$ in (3·1·3A) is defined as the u^{th} row of the N x k matrix $\underset{\sim}{V}$, then from (3·1·3A),

$$\begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1k} \\ \vdots & \vdots & & \vdots \\ v_{N1} & v_{N2} & \cdots & v_{Nk} \end{bmatrix} = \underset{\sim}{X} \begin{bmatrix} \frac{1}{h_1} & & \underset{\sim}{0} \\ & \ddots & \\ \underset{\sim}{0} & & \frac{1}{h_k} \end{bmatrix}. \qquad (3.1.3C)$$

From the relation (3.1.3C), the general linear model in the intermediate variables can now be written as

$$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{X}\,\underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} \\ &= \underset{\sim}{V}\,\underset{\sim}{H}\,\underset{\sim}{\gamma} + \underset{\sim}{\varepsilon} \\ &= \underset{\sim}{V}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \qquad (3.1.4) \end{aligned}$$

where $\underset{\sim}{\beta}=\underset{\sim}{H}\underset{\sim}{\gamma}$ and $\underset{\sim}{\gamma}$ is the vector of unknown parameters in the model (3.0.1).

Since the transformation (3.1.1) changes only the scale of the variables, the rank of the matrix $\underset{\sim}{V}$ in (3.1.4) is k-q also. We could consider a reduced problem involving only k-q components by ignoring one of the components in each category. In Section 2.2, it was explained why this is undesirable, namely that the effects of the ignored components are obscured. Thus, instead of ignoring components (and at the same time ignoring intermediate variables), we shall reparametrize the model (3.1.4) to one of full rank. By a reparametrization of the model is meant making a transformation from the vector $\underset{\sim}{\beta}$ to a vector $\underset{\sim}{\alpha}$ in the new model $\underset{\sim}{y} = \underset{\sim}{W}\,\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$ so that each element of the vector is estimable. Then an estimate of each β_j $(1\leq j\leq k)$ in (3.1.4) and therefore an estimate of each γ_j $(1\leq j\leq k)$ in (3.0.1) can be obtained and we shall force these estimates to be unique as we show

later in the next section.

3.2 Transformation to independent variables.

We now make a transformation from the system of the intermediate variables v_j $(1 \le j \le k)$ to a system involving new variables w_i $(1 \le i \le k)$. The transformation is such that every point in the (k-q)-dimensional unit spherical region has the value zero for the last q of the w_i's, that is, $w_i=0$ for $k-q+1 \le i \le k$. This in effect reduces the problem of k variables to the k-q variables w_i $(1 \le i \le k-q)$ which are mathematically independent.

The constraint $x_1+x_2+\ldots+x_{n_1}=\frac{1}{q}$ gives rise to the homogeneous property $h_1v_1+h_2v_2+\ldots+h_{n_1}v_{n_1}=0$ of the intermediate variables which defines a (k-1)-dimensional linear manifold in k-space. The q constraints $h_av_a+\ldots+h_{a+1}v_{a+1}=0$ where $a=S_i$ and $S_i=n_0+n_1+\ldots+n_i$ $(0 \le i \le q-1)$ define a (k-q)-dimensional linear manifold. The transformation which we use is a rotation of the axes of the intermediate variables about the origin $\underset{\sim}{v}=\underset{\sim}{0}$. The rotation enables us to express the constraints in the form $w_i=0$ $(k-q+1 \le i \le k)$. Then by ignoring these zero coordinates, we in effect project the k-dimensional unit sphere onto the (k-q)-dimensional manifold and it becomes a (k-q)-dimensional sphere which is again a unit sphere since the manifold, specified by the constraints, passes through the origin. Therefore, the region of interest will now be a (k-q)-dimensional unit spherical region and will be

centered at $\underset{\sim}{w}=\underset{\sim}{0}$ where $\underset{\sim}{w}'=(w_1,w_2,\ldots,w_{k-q},0,0,\ldots,0)$ or $\underset{\sim}{w}'=(w_1,w_2,\ldots,w_{k-q})$. (We shall consider both forms of the vector $\underset{\sim}{w}'$ to be equivalent.)

Since the rank of the N x k matrix $\underset{\sim}{V}$ is k-q, there are q independent linear relations among the k columns. Suppose the transformation of the variables v_j $(1\leq j\leq k)$ is defined by

$$\underset{\sim}{V}\ \underset{\sim}{T} = [\tilde{\underset{\sim}{W}}\ \underset{\sim}{0}] \qquad (3.2.1)$$

where $\tilde{\underset{\sim}{W}}$ is an N x (k-q) matrix, $\underset{\sim}{0}$ is an N x q matrix of zeroes and $\underset{\sim}{T}$ is a k x k non-singular matrix. Since rank is unaffected by the multiplication of a non-singular matrix, the rank of the matrix $\tilde{\underset{\sim}{W}}$ is k-q. If we also specify that the matrix $\underset{\sim}{T}$ be orthogonal and in addition, partition $\underset{\sim}{T}$ by letting

$$\underset{\sim}{T} = [\underset{\sim}{T}_1\ \underset{\sim}{T}_2] \qquad (3.2.2)$$

where $\underset{\sim}{T}_1$ is k x (k-q) and $\underset{\sim}{T}_2$ is k x q, then from (3.2.1), we see that the matrix $\underset{\sim}{T}_2$ is defined such that $\underset{\sim}{V}\ \underset{\sim}{T}_2 = \underset{\sim}{0}$.

Let the matrix $\underset{\sim}{T}_1$ be <u>any</u> arbitrary k x (k-q) matrix containing (k-q) orthonormal columns such that $\underset{\sim}{T}$ in (3.2.2) is orthogonal. The arbitrariness of the elements of the matrix $\underset{\sim}{T}_1$ is justified at the beginning of Chapter 5 when a special class of designs is suggested for use.

The model (3.1.4) can now be written as

$$\begin{aligned}\underset{\sim}{y} &= \underset{\sim}{V}\ \underset{\sim}{T}\ \underset{\sim}{T}'\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon} \\ &= \underset{\sim}{V}\ [\underset{\sim}{T}_1\ \underset{\sim}{T}_2] \begin{bmatrix}\underset{\sim}{T}_1' \\ \underset{\sim}{T}_2'\end{bmatrix} \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}\end{aligned}$$

and since $\underset{\sim}{V}\ \underset{\sim}{T}_2 = \underset{\sim}{0}$, we see that

$$\underset{\sim}{y} = \underset{\sim}{V}\ \underset{\sim}{T}_1\ \underset{\sim}{T}_1'\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}. \tag{3.2.3}$$

Let

$$\underset{\sim}{V}\ \underset{\sim}{T}_1 = \underset{\sim}{\tilde{W}} \tag{3.2.4}$$

and

$$\underset{\sim}{T}_1'\ \underset{\sim}{\beta} = \underset{\sim}{\alpha}_3. \tag{3.2.5}$$

The model (3.2.3) can now be written as

$$\underset{\sim}{y} = \underset{\sim}{\tilde{W}}\ \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon} \tag{3.2.6}$$

where $\underset{\sim}{\tilde{W}}$ is the N x (k-q) matrix of rank k-q in (3.2.1) and (3.2.4) and we shall call $\underset{\sim}{\alpha}_3$ the vector of coefficients associated with the reduced model (3.2.6) to distinguish it from $\underset{\sim}{\alpha}_1$, $\underset{\sim}{\alpha}_2$ introduced later in Chapter 4.

To review the reparametrization of the model as well as to show one construction of the matrix $\underset{\sim}{T}$ in (3.2.2), let q=2 and designate $x_1, x_2, \ldots, x_n$ as belonging to category 1 and $x_{n+1}, \ldots, x_k$ as being from category 2. Then from (1.3.2),

$$\sum_{i=1}^{n} x_i = \frac{1}{2}\ ,\quad \sum_{j=n+1}^{k} x_j = \frac{1}{2}$$

and the same constraints are placed on the x_{0i} and x_{0j}, i.e.,

$$\sum_{i=1}^{n} x_{0i} = \sum_{i=1}^{n} \sum_{u=1}^{N} \frac{x_{ui}}{N}$$

$$= \sum_{u=1}^{N} \frac{\frac{1}{2}}{N} = \sum_{j=n+1}^{k} x_{0j}. \tag{3.2.7}$$

We have adopted the convention that the central value x_{0j} is

the mean of the x_{uj} $(1\le u\le N, 1\le j\le k)$ and therefore, the origin is the mean of the v_j $(1\le j\le k)$ so that $\sum_{u=1}^{N} v_{ui}=0$. Now since

$$v_i = \frac{x_i - x_{0i}}{h_i} \qquad (1\le i\le k),$$

then

$$\sum_{i=1}^{n} h_i v_i = \sum_{i=1}^{n} (x_i - x_{0i})$$

$$= 0$$

and

$$\sum_{i=1}^{n} h_i v_i = \sum_{j=n+1}^{k} h_j v_j = \sum_{j=n+1}^{k} (x_j - x_{0j})$$

$$= 0. \qquad (3.2.8)$$

To determine a form of the matrix $\underset{\sim}{T}_2$, let us normalize the h_i and h_j $(1\le i\le n,\ n+1\le j\le k)$ in (3.2.8) by defining

$$\tilde{h}_i = \frac{h_i}{\left(\sum_{i=1}^{n} h_i^2\right)^{1/2}} \qquad i=1,2,\ldots,n$$

$$\tilde{h}_j = \frac{h_j}{\left(\sum_{j=n+1}^{k} h_j^2\right)^{1/2}} \qquad j=n+1,\ldots,k. \qquad (3.2.9)$$

Then (3.2.8) can be rewritten as

$$\sum_{i=1}^{n} \tilde{h}_i v_i = \sum_{j=n+1}^{k} \tilde{h}_j v_j = 0. \qquad (3.2.10)$$

If we now define the vectors $\underset{\sim}{\tilde{h}}'$ and $\underset{\sim}{\bar{h}}'$ by

$$\tilde{\underset{\sim}{h}}' = (\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_n, 0, 0, \ldots, 0),$$
$$\bar{\underset{\sim}{h}}' = (0, 0, \ldots, 0, \tilde{h}_{n+1}, \tilde{h}_{n+2}, \ldots, \tilde{h}_k), \qquad (3.2.11)$$

where $\tilde{\underset{\sim}{h}}'$ and $\bar{\underset{\sim}{h}}'$ are both 1 x k, then the matrix $\underset{\sim}{T}_2$ in (3.2.2) can be defined by

$$\underset{\sim}{T}_2 = [\ \tilde{\underset{\sim}{h}}\ \bar{\underset{\sim}{h}}\ ], \qquad (3.2.12)$$

so that with the form of the matrix $\underset{\sim}{V}$ defined as in (3.1.3C), $\underset{\sim}{V}\ \underset{\sim}{T}_2 = \underset{\sim}{0}$.

We can also use the h_i $(1 \le i \le k)$ to define the elements of the k x (k-2) matrix $\underset{\sim}{T}_1$ so that the matrix $[\underset{\sim}{T}_1\ \underset{\sim}{T}_2]$ is orthogonal where $\underset{\sim}{T}_2$ is defined by (3.2.12). For example, if we denote by t_{ij} the element in the i^{th} row and j^{th} column of the matrix $\underset{\sim}{T}_1$, then by letting

$$t_{11} = h_2 \qquad t_{ij} = h_i h_{j+1} \quad \begin{matrix} j=2,3,\ldots,n-1 \\ i=1,2,\ldots,j \end{matrix}$$

$$t_{21} = -h_1 \qquad t_{ij} = -\left(\sum_{\ell=1}^{j} h_\ell^2\right) \quad \begin{matrix} j=2,3,\ldots,n-1 \\ i=j+1 \end{matrix}$$

$$t_{ij} = 0 \quad \begin{matrix} j=1,2,\ldots,n-2 \\ i=j+2,\ldots,n \end{matrix} \qquad t_{ij} = 0 \quad \begin{matrix} i=1,2,\ldots,n \\ j=n,\ldots,k-2 \end{matrix} \qquad (3.2.13)$$

$$t_{n+1,n} = h_{n+2} \qquad t_{ij} = h_i h_{j+1} \quad \begin{matrix} j=n+2,\ldots,k-1 \\ i=1,2,\ldots,j \end{matrix}$$

$$t_{n+2,n} = -h_{n+1} \qquad t_{ij} = -\left(\sum_{\ell=n+1}^{j+1} h_\ell^2\right) \quad \begin{matrix} j=n+1,\ldots,k-2 \\ i=j+2,\ldots,k \end{matrix}$$

$$t_{ij} = 0 \quad \begin{matrix} j=n,\ldots,k-3 \\ i=j+2,\ldots,k \end{matrix} \qquad t_{ij} = 0 \quad \begin{matrix} i=1,2,\ldots,n \\ j=n,\ldots,k-2 \end{matrix}$$

and column-normalizing the above elements, the matrix $\underset{\sim}{T}$ is orthogonal. For example, let n=6 and k=10. The matrix $\underset{\sim}{T}$, before column-normalizing the elements of the matrix $\underset{\sim}{T}_1$, is

$$
\begin{bmatrix}
h_2 & h_1h_3 & h_1h_4 & h_1h_5 & h_1h_6 & & & & \tilde{h}_1 & 0 \\
-h_1 & h_2h_3 & h_2h_4 & h_2h_5 & h_2h_6 & & & & \tilde{h}_2 & 0 \\
 & -\left(\sum_{\ell=1}^{2} h_\ell^2\right) & h_3h_4 & h_3h_5 & h_3h_6 & & \underset{\sim}{0} & & \tilde{h}_3 & 0 \\
 & & -\left(\sum_{\ell=1}^{3} h_\ell^2\right) & h_4h_5 & h_4h_6 & & & & \tilde{h}_4 & 0 \\
 & & & -\left(\sum_{\ell=1}^{4} h_\ell^2\right) & h_5h_6 & & & & \tilde{h}_5 & 0 \\
 & & & & \left(\sum_{\ell=1}^{5} h_\ell^2\right) & & & & \tilde{h}_6 & 0 \\
 & & & & & h_8 & h_7h_9 & h_7h_{10} & 0 & \tilde{h}_7 \\
 & & & & & -h_7 & h_8h_9 & h_8h_{10} & 0 & \tilde{h}_8 \\
 & & \underset{\sim}{0} & & & & \left(\sum_{\ell=7}^{8} h_\ell^2\right) & h_9h_{10} & 0 & \tilde{h}_9 \\
 & & & & & & & \left(\sum_{\ell=7}^{9} h_\ell^2\right) & 0 & \tilde{h}_{10}
\end{bmatrix}
$$

We could also use the Gram-Schmidt orthogonalization process (see Appendix F) for the construction of the matrix $\underset{\sim}{T}_1$, but the above construction appears to be more expedient.

Let us now summarize the last two sections. The mixture components were transformed and the general linear model redefined in terms of the intermediate variables by

$$\underset{\sim}{y} = \underset{\sim}{V}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}. \qquad (1)$$

Since the matrix $\underset{\sim}{V}$ in (1) is less than full rank, a

reparametrization of the model (1) to one of full rank in new variables w_i $(1 \le i \le k-q)$, was performed and the general linear model rewritten as

$$\underset{\sim}{y} = \tilde{\underset{\sim}{W}}\ \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon} \ . \qquad (2)$$

The variables w_i $(1 \le i \le k-q)$ are called <u>design variables</u> because we shall construct designs (Chapter 5 and 7) in the $(k-q)$-coordinate system of the w_i. By using an orthogonal matrix in the reparametrization of the model, distances in both the $\underset{\sim}{v}$ metric and $\underset{\sim}{w}$ metric remained preserved and equal. This enables us to work with the design variables w_i $(1 \le i \le k-q)$ for the purpose of design construction, etc., requiring only a simple inverse transformation to redefine the problem in the intermediate variables v_j $(1 \le j \le k)$.

To estimate the response with a model in the intermediate variables v_j, we see from equation (3.2.5) that the β_j $(1 \le j \le k)$ are related to the α_i $(1 \le i \le k-q)$, the coefficients of the design variables by

$$\underset{\sim}{T}_1' \ \underset{\sim}{\beta} = \underset{\sim}{\alpha}_3 \ . \qquad (3)$$

In order to obtain the estimates $\hat{\underset{\sim}{\beta}}$ of $\underset{\sim}{\beta}$ in (3), constraints must be placed on the estimates $\hat{\underset{\sim}{\beta}}$. This is seen from the model

$$\underset{\sim}{y} = \underset{\sim}{V}\ \underset{\sim}{T}\ \underset{\sim}{T}'\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon} = [\tilde{\underset{\sim}{W}}\ \underset{\sim}{Q}]\ \underset{\sim}{T}'\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$$
$$= [\tilde{\underset{\sim}{W}}\ \underset{\sim}{Q}]\ \underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$$

where $\underset{\sim}{\alpha}' = (\underset{\sim}{\alpha}_3', \underset{\sim}{\alpha}_{(rest)}')$, and $\underset{\sim}{T}'\underset{\sim}{\beta} = \underset{\sim}{\alpha}$, i.e.,

$$\underset{\sim}{\beta} = \underset{\sim}{T}_1 \underset{\sim}{\alpha}_3 + \underset{\sim}{T}_2 \underset{\sim}{\alpha}_{(rest)} \cdot \qquad (4)$$

Since the last q columns of the matrix $[\underset{\sim}{\tilde{W}} \; \underset{\sim}{0}]$ contain all zeroes, the effect of multiplying this matrix by $\underset{\sim}{\alpha}$ is to annihilate the elements of $\underset{\sim}{\alpha}_{(rest)}$ whose components are, for the time being, of no interest to us. However, putting $\underset{\sim}{\alpha}_{(rest)} = \underset{\sim}{0}$ and letting $\hat{\underset{\sim}{\alpha}}_3$ be the estimate of $\underset{\sim}{\alpha}_3$ in (4), the estimates $\hat{\underset{\sim}{\beta}}$ can be calculated by setting $\hat{\underset{\sim}{\beta}} = \underset{\sim}{T}_1 \hat{\underset{\sim}{\alpha}}_3$ and $\underset{\sim}{T}_2' \hat{\underset{\sim}{\beta}} = \underset{\sim}{0}$ from (4). That is to say, by fixing $\underset{\sim}{\alpha}_{(rest)} = \underset{\sim}{0}$, we force the estimates $\hat{\underset{\sim}{\beta}}$ in (3) to be unique and the most convenient way of doing this was by setting $\underset{\sim}{\alpha}_{(rest)} = \underset{\sim}{0}$. (Towards the end of the present section, we shall indicate that there is, after all, some point in taking $\underset{\sim}{\alpha}_{(rest)} \neq \underset{\sim}{0}$.)

If the number of design variables (k-q) is small, it may be to our advantage to use the model in the design variables w_i for the purposes of prediction, etc. This is so for two reasons; the model (2) contains q fewer variables than the model (1) and the estimates of the parameters in (2) are not subject to any constraints. For example, if k=4, q=2 and

$$\underset{\sim}{T}_1 = \begin{bmatrix} .707 & 0 \\ -.707 & 0 \\ 0 & .707 \\ 0 & -.707 \end{bmatrix},$$

then

$$\hat{\underset{\sim}{\beta}} = \begin{bmatrix} \hat{\beta}_1 \\ -\hat{\beta}_1 \\ \hat{\beta}_3 \\ -\hat{\beta}_3 \end{bmatrix}$$

and the prediction models in the intermediate variables and in the design variables respectively are

$$\hat{y} = \hat{\beta}_1 v_1 - \hat{\beta}_1 v_2 + \hat{\beta}_3 v_3 - \hat{\beta}_3 v_4 ,$$

$$\hat{y} = \hat{\alpha}_1 w_1 + \hat{\alpha}_2 w_2 .$$

In the second model, $\hat{\alpha}_i$ is a measure of the slope of the response surface in the direction of the i^{th} $(1 \le i \le k-q)$ coordinate axis. In the first model however, the meaning of $\hat{\beta}_i$ is not clear.

In order to determine the experimental settings of the mixture components, we know that in the model $\underset{\sim}{y} = \tilde{\underset{\sim}{W}}\, \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}$, the matrix $\tilde{\underset{\sim}{W}}$ contains the levels of the design variables. That is to say, $\tilde{\underset{\sim}{W}}$ corresponds to the matrix $\underset{\sim}{D}_w$ mentioned in Section 1.4. However, although the experimenter designates the levels of the design variables to be run in the N experiments, the actual experimentation involves the mixture components. Therefore to compute the levels of the mixture components to be performed experimentally, we have from (3.2.4) and (3.1.3),

$$[\tilde{\underset{\sim}{W}}\ \underset{\sim}{0}] = \underset{\sim}{V}\ \underset{\sim}{T}$$

$$= \underset{\sim}{X}\ \underset{\sim}{H}^{-1}\ \underset{\sim}{T} \qquad (3.2.14)$$

where $\underset{\sim}{0}$ is an N x q matrix of zeroes and the elements of the matrix $\underset{\sim}{X}$ are shown in (3·1·3B). Then from (3·2·14),

$$\underset{\sim}{X} = [\tilde{\underset{\sim}{W}}\ \underset{\sim}{0}] \begin{bmatrix} \underset{\sim}{T}_1' \\ \underset{\sim}{T}_2' \end{bmatrix} \quad \underset{\sim}{H}$$

$$= \tilde{\underset{\sim}{W}}\ \underset{\sim}{T}_1'\ \underset{\sim}{H} \ . \qquad (3\cdot2\cdot14A)$$

If we denote the elements of $\tilde{\underset{\sim}{W}}$ by w_{uj} $(1 \le u \le N,\ 1 \le j \le k-q)$, and the elements of $\underset{\sim}{T}_1'$ by $t_{j\ell}$ $(1 \le \ell \le k-q,\ 1 \le j \le k)$, that is,

$$\tilde{\underset{\sim}{W}} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1k-q} \\ w_{21} & w_{22} & \cdots & w_{2k-q} \\ & & & \\ w_{N1} & w_{N2} & \cdots & w_{Nk-q} \end{bmatrix}, \quad \underset{\sim}{T}_1' = \begin{bmatrix} t_{11} & t_{21} & \cdots & t_{k1} \\ t_{12} & t_{22} & \cdots & t_{k2} \\ & & & \\ t_{1k-q} & t_{2k-q} & \cdots & t_{kk-q} \end{bmatrix}$$

then from (3·2·14A),

$$x_{11} - x_{01} = (w_{11}t_{11} + w_{12}t_{12} + \cdots + w_{1k-q}t_{1k-q})\ h_1 \ ,$$

$$x_{12} - x_{02} = (w_{11}t_{21} + w_{12}t_{22} + \cdots + w_{1k-q}t_{2k-q})\ h_2$$

$$x_{21} - x_{01} = (w_{21}t_{11} + w_{22}t_{12} + \cdots + w_{2k-q}t_{1k-q})\ h_1$$

or

$$x_{uj} = \Big(\sum_{\ell=1}^{k-q} w_{u\ell}t_{j\ell}\Big)\ h_j + x_{0j} \ , \qquad (3\cdot2\cdot15)$$

where $1 \le u \le N$, $1 \le j \le k$. The equation (3·2·15) will be used for calculating the level of the j^{th} $(1 \le j \le k)$ mixture component to be used in the u^{th} $(1 \le u \le N)$ experiment.

Up to now we have considered only the first-degree model for fitting the response surface. The reason for this is that this model is usually employed at the initial stage of experimentation. Also linear models are sometimes considered to be adequate when the ranges of the variables are not very wide. However, since the emphasis in this paper is on using response surface methods, and the usual linear model representation contains a constant term, for example,

$$y = \alpha_0 + \sum_{i=1}^{k-q} \alpha_i w_i + \epsilon ,$$

we shall include the constant term in the model in the future.

In view of the fact that there are q columns containing all zeroes in the matrix $[\tilde{W}\ \underline{0}]$, we may, by placing non-zero elements in these q columns, extract information from the observations, of a completely different nature from that of the design variables. Hence, we next propose to change the linear model (3.2.6) in the design variables to a linear model consisting not only of the design variables but additional variables which we shall call "block-associated" variables. (The name "block-associated" is used because the concept is identical to the use of block variables.)

The design variables w_i $(1 \leq i \leq k-q)$ are linear combinations of the mixture components although their physical dimensions are eliminated by the h_i $(1 \leq i \leq k)$. The block-associated variables are not linear combinations of the

chemical components. Also, the block-associated variables are not merely of different physical nature from the w_i $(1 \leq i \leq k-q)$, but are of different mathematical nature in that they are discrete rather than continuous. Therefore, the block-associated variables will not be denoted by w_j where $j=k-q+1,\ldots,k$ which would be misleading but instead will be denoted by z_i $(1 \leq i \leq q)$. (The inclusion of the block-associated variables will not affect the position of the design points in the design space whose coordinates are at the $w_1, w_2, \cdots, w_{k-q}$.)

In the original model, (3.2 6), we assumed no blocking of the observations and the block-associated variables would have served no purpose. When there is blocking, that is, when the complete set of N experiments is performed in stages and there might be variation in chemical or physical conditions from stage to stage, then it is useful to make use of this revised model. The number of block-associated variables we add can be any positive integer ν such that $k-q+\nu \leq N$, although the original matrix $\underset{\sim}{Q}$ in (3.2.6) has only q columns containing all zeroes. If ν is greater than q, we have to add elements to $\underset{\sim}{\alpha}_{(rest)}$ since $\underset{\sim}{\alpha}_{(rest)}$ contains only q elements. We have however decided, admittedly somewhat arbitrarily, that in this paper, the number ν will not exceed q.

In the next section, we replace the columns containing all zeroes in the matrix $[\underset{\sim}{\tilde{W}}\ \underset{\sim}{Q}]$ with one column containing N ones and q-1 columns whose elements we show to be the

coefficients in q-1 linear functions of the observations. We shall call this addition of columns of non-zero elements, "augmenting" the matrix $\tilde{\underset{\sim}{W}}$. The augmented matrix will be called $\underset{\sim}{W}$. It is shown in Section 3.4, that by constructing the N x k' (where k'=k-q+ν) matrix $\underset{\sim}{W}$ to be of rank k', the estimates $\hat{\underset{\sim}{\alpha}}_3$ of the effects of the design variables will be unaffected by the addition of the ν terms to the model.

3.3 Augmenting the matrix of design variables: Orthogonal contrast approach.

In this section, we are faced with the problem of how to construct the N x q matrix $\underset{\sim}{C}$ (where now ν=q) in

$$\underset{\sim}{W} = [\tilde{\underset{\sim}{W}}\ \underset{\sim}{C}] \qquad (3.3.0)$$

so that the matrix $\underset{\sim}{W}$ is of full rank. One method certainly is by finding an N x q matrix $\underset{\sim}{C}$ whose columns are orthogonal to the columns of $\tilde{\underset{\sim}{W}}$. A matrix $\underset{\sim}{C}_1$ will be used later in a slightly different sense.

Now we have stated that the elements of the matrix $\tilde{\underset{\sim}{W}}$ are the levels of the design variables to be used in the N experiments. To be more explicit, w_{ui} ($1 \le u \le N$, $1 \le i \le k-q$) is the element in the u^{th} row and i^{th} column of the matrix $\tilde{\underset{\sim}{W}}$ and is the level of the i^{th} design variable set by the experimenter to be used in the u^{th} experiment. By knowing the types of designs (rotatable designs) that we shall be using and therefore knowing the general form of the matrix $\tilde{\underset{\sim}{W}}$, we are now in a position to derive the elements in the q columns

of the matrix $\underset{\sim}{C}$.

One method of adding terms to the model is derived by Box and Hunter [6] while discussing orthogonal blocking. We shall use another method which involves somewhat simpler computations in which the contrasts are orthogonal over the whole experiment. This involves the inversion of a diagonal rather than a slightly more complicated matrix when obtaining the estimates of the parameters in the model (see Chapter 6).

The theory behind orthogonal contrasts as well as orthogonal blocking is the placing of patterns of experimental design points in row partitions (compartments) of the design matrix so that each compartment contains a first-degree orthogonal design. By row partitioning is meant dividing the design matrix in such a way that each compartment contains some number of rows of the matrix. A first-degree orthogonal design is a design such that the columns of the design matrix are orthogonal to one another. For example, if $\underset{\sim}{w}_i$ $(1 \leq i \leq k-q)$ is a column of the matrix $\underset{\sim}{\tilde{W}}$, then $\underset{\sim}{w}_i' \underset{\sim}{w}_j = 0$ for all $1 \leq i \neq j \leq k-q$. In this way then, if the execution of the entire program of N experiments is performed by compartments and the compartments are not too large, the experiments in each compartment of the design matrix might well be performed under somewhat homogeneous environmental conditions.

To develop the theory involving orthogonal contrasts, let $z_{i1}, z_{i2}, \ldots, z_{iN}$ be any set of coefficients not all

zero such that

$$\sum_{u=1}^{N} z_{iu} = 0 \qquad (3.3.1)$$

where i = 1, 2, 3, ... corresponds to the contrasts 1, 2, 3, ... If we also define

$$L_i = \sum_{u=1}^{N} z_{iu}\, y_u \,, \qquad (3.3.2)$$

then L_i is said to be a contrast (the i^{th} contrast) of the N observations. Let the coefficients z_{1u}, z_{2u} (u = 1, 2, ... N) correspond to two such contrasts L_1 and L_2. Then the contrasts L_1 and L_2 are said to be orthogonal if

$$\sum_{u=1}^{N} z_{1u} z_{2u} = 0. \qquad (3.3.3)$$

The equalities (3.3.1) and (3.3.3) show that the properties of being a contrast and of orthogonality are independent of the particular y_u and depend only on the coefficients z_{iu}.

In constructing orthogonal contrasts, a multiplicity of different methods (except when N=2) can be used to form the z_{iu}. One method which can easily be verified is to let

$$\begin{aligned} z_{iu} &= 1 && \text{if } 1 \le u \le i \\ &= -i && \text{if } u = i + 1 \\ &= 0 && \text{if } i + 1 < u \le N. \end{aligned} \qquad (3.3.4)$$

If we have i = 1, 2, 3, the matrix of coefficients using

(3.3.4) is

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 1 & 1 \\ 0 & -2 & 1 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \\ \vdots & \vdots & \vdots \\ 0 & 0 & 0 \end{bmatrix} .$$

Suppose now that the concentration is on groups or blocks of observations where each contrast is a contrast between $\underline{a}$ groups. Note, in this paper, the word "group" is never used in the sense of the theory of groups. Let the total of N_j observations corresponding to the j^{th} group be denoted by Y_j (for example, $Y_1 = y_1 + y_2 + \ldots + y_{N_1}$) and such that $N_1 + N_2 + \ldots + N_a = N$. A lemma on orthogonal contrasts is stated [11] approximately in the following manner. The first (a-1) orthogonal contrasts can be so defined that for each i and j ($1 \leq i \leq a-1$, $1 \leq j \leq a$), if z_{iu} is independent of u for all u in the j^{th} group, the contrasts are then contrasts between the quantities $\frac{Y_j}{N_j}$, the means of the $\underline{a}$ groups. Moreover (as is shown in some textbooks on experimental design, see for example, Cochran and Cox, second edition, pages 62-67),

$$\sum_{j=1}^{a} \frac{Y_j^2}{N_j} - \frac{\left(\sum_{u=1}^{N} y_u\right)^2}{N} = \sum_{i=1}^{a-1} \frac{L_i^2}{\sum_{u=1}^{N} z_{iu}^2} \qquad (3.3.5)$$

is the sum of squares between the means of the $\underline{a}$ groups. In this way then, the idea of orthogonal contrasts of the observations may be developed, having the property that any contrast between the means of the groups is orthogonal with any other contrast between the means.

With this concept of "grouped orthogonal contrasts", if we let $\underline{a} = q$, we can set up $q - 1$ orthogonal contrasts where each contrast will be involved with at least q blocks or groups of observations. The coefficients z_{iu} $(1 \leq i \leq q-1,\ 1 \leq u \leq N)$ will be the elements in $q - 1$ columns added to the matrix $\tilde{\underset{\sim}{W}}$. The theory in this paper will be developed so that for first-degree designs, the number of groups of observations (compartments in $\underset{\sim}{D}_w$) will be equal to q. For second-degree designs, the number of groups of observations will be greater than or equal to q, where, again, q is the number of categories of mixture components.

To determine the value of z_{iu} (the u^{th} coefficient in the i^{th} contrast), let N_j be the number of rows of $\tilde{\underset{\sim}{W}}$ which are contained in the j^{th} compartment $(1 \leq j \leq q)$. Since the z_{iu} are equal for all rows in a particular compartment for each contrast, we can denote the coefficients in the j^{th} compartment corresponding to the i^{th} contrast by $z_{i.}^{(j)}$. Now, let

$$z_{i.}^{(j)} = 1 \qquad 1 \leq j \leq i\ , \qquad (3.3.6)$$

and

$$z_{i.}^{(j)} = 0 \qquad i + 2 \leq j \leq q. \qquad (3.3.6A)$$

(The value $z_{i.}^{(i+1)}$ will be derived shortly.) Then

$$\sum_{u=1}^{N} z_{iu} = 0$$

$$= \sum_{j=1}^{q} \sum_{u=N_{j-1}+1}^{N_j} z_{iu}$$

$$= \sum_{j=1}^{q} N_j z_{i.}^{(j)}$$

$$= \sum_{j=1}^{i} N_j + N_{i+1} z_{i.}^{(i+1)} + \sum_{j'=i+2}^{q} N_{j'} z_{i.}^{(j')} \qquad (3.3.6B)$$

From (3.3.6A),

$$N_{i+1} z_{i.}^{(i+1)} = - \sum_{j=1}^{i} N_j$$

and therefore,

$$z_{i.}^{(i+1)} = - \frac{\sum_{j=1}^{i} N_j}{N_{i+1}} \qquad (3.3.7)$$

Equation (3.3.7) shows how to calculate the value of the N_{i+1} $(1 \leq i \leq q-1)$ coefficients of the i^{th} contrast which correspond to the rows of the $i+1^{st}$ compartment of the matrix $\underset{\sim}{W}$. When the N_j $(1 \leq j \leq q)$ are equal in number, equation (3.3.7) is

$$z_{i.}^{(i+1)} = -i, \qquad (3.3.8)$$

and we are reduced to the case of equations (3.3.4).

The elements in the last q-1 columns of the matrix $\underset{\sim}{W}$ are now defined by equations (3.3.6), (3.3.6A) and (3.3.7) for i = 1, 2, ..., q-1. It is easy to see that if in fact the execution of the N experiments is performed according to the compartments of the matrix $\underset{\sim}{W}$, the information obtained from the contrast terms in the model is similiar to that which would have been obtained had we used the method of orthogonal blocking. For instance, with N_j $(1 \leq j \leq q)$ defined as the number of rows in the j^{th} compartment of the matrix $\underset{\sim}{W}$, the estimate $\hat{\alpha}_{L_i}$ of the parameter α_{L_i} associated with the i^{th} $(1 \leq i \leq q-1)$ contrast term in the model is

$$\hat{\alpha}_{L_i} = \frac{\dfrac{1}{\sum_{j=1}^{i} N_j} \sum_{j=1}^{i} N_j \bar{y}_j - \bar{y}_{i+1}}{(N_{i+1} + \sum_{j=1}^{i} N_j) / N_{i+1}}, \quad (1 \leq i \leq q-1) \qquad (3.3.9)$$

where $\bar{y}_j$ is the mean of the j^{th} group of observations. The equation (3.3.9) was found by the formula

$$(\underset{\sim}{W}'\underset{\sim}{W})^{-1} \underset{\sim}{W}'\underset{\sim}{y} = \frac{L_i}{\underset{\sim}{w}_i'\underset{\sim}{w}_i}$$

where $\underset{\sim}{w}_i$ is the column of the matrix $\underset{\sim}{W}$ which contains the coefficients z_{iu} $(1 \leq u \leq N)$ of the contrast L_i. When the N_j

are equal in size, equation (3.3.9) is simplified to

$$\hat{\alpha}_{L_i} = \frac{\sum_{j=1}^{i} \bar{y}_j - i\bar{y}_{i+1}}{i(i+1)} \qquad (1 \leq i \leq q-1), \qquad (3.3.10)$$

where $\bar{y}_j$ is again the mean of the j^{th} group of observations.

Since we now have a method for adding q-1 additional columns to the matrix $\underset{\sim}{W}$, the final column in $\underset{\sim}{W}$ needed to make the rank of the $\underset{\sim}{W}$ matrix equal to k will be a column containing all ones. The usual terminology in response surface work is to call this column the "mean" column since it corresponds to the constant term in the model and the least squares estimate of the constant term is

$$\bar{y} = \frac{1}{N} \sum_{u=1}^{N} y_u.$$

In the addition of the q terms to the model, we stated that the estimates of the effects of the design variables are unaffected if the addition is achieved with an N x q matrix $\underset{\sim}{C}$ whose columns are orthogonal to the columns of the matrix $\underset{\sim}{\tilde{W}}$. Since we have just derived the elements of the q columns of the matrix $\underset{\sim}{C}$ (one column containing N ones and q-1 columns containing the contrast coefficients), we now show that the estimates of the effects of the design variables are indeed unaffected by the addition.

3.4 The effect of the augmentation of the matrix $\tilde{W}$ on the estimates of the effects of the design variables.

Let the matrix $\underset{\sim}{W}$ be defined by

$$\underset{\sim}{W} = [\hat{\underset{\sim}{W}} \; \underset{\sim}{C}_1] \tag{3.4.1}$$

where

$$\hat{\underset{\sim}{W}} = [\underset{\sim}{j}_N \; \tilde{\underset{\sim}{W}}]$$

and $\underset{\sim}{C}_1$ is an N x (q-1) matrix of contrast coefficients. Note that $\nu = q$ and the columns of the matrix $\underset{\sim}{W}$ have been rearranged placing $\underset{\sim}{j}_N$ as the first column. This was done so that the mean term is the first term in the model. This is the usual form of the model. Actually, we did not have to rearrange the columns of $\underset{\sim}{W}$ but rather rearrange the columns of the matrix $\underset{\sim}{T}$ so that when q=2,

$$\underset{\sim}{T} = [\tilde{\underset{\sim}{h}} \; \underset{\sim}{T}_1 \; \tilde{\underset{\sim}{h}}] \; .$$

Then the transformation from the variables v_j to the design variables w_i $(1 \leq i \leq k-q)$ is

$$\underset{\sim}{V}\underset{\sim}{T} = [\underset{\sim}{0} \; \tilde{\underset{\sim}{W}} \; \underset{\sim}{0}]$$

where $\underset{\sim}{0}$ is an N x 1 column of zeroes and the addition of non-zero elements to the columns of zeroes results in the matrix $\underset{\sim}{W}$ in (3.4.1).

The general linear model can now be written as

$$\begin{aligned} \underset{\sim}{y} &= \underset{\sim}{W} \underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} \\ &= [\hat{\underset{\sim}{W}} \; \underset{\sim}{C}_1] \begin{bmatrix} \underset{\sim}{\alpha}_1 \\ \underset{\sim}{\alpha}_4 \end{bmatrix} + \underset{\sim}{\varepsilon} \end{aligned} \tag{3.4.2}$$

where $\underset{\sim}{\alpha}_1' = (\alpha_0, \underset{\sim}{\alpha}_3')$ and $\underset{\sim}{\alpha}_4' = (\alpha_{k-q+1}, \alpha_{k-q+2}, \ldots, \alpha_k)$ $= (\alpha_{L_1}, \alpha_{L_2}, \ldots, \alpha_{L_{q-1}})$. If we put p=k-q, then (3.4.2) can be written as

$$\underset{\sim}{y} = [\underset{\sim}{j}_N \; \underset{\sim}{\tilde{W}} \; \underset{\sim}{C}_1] \begin{bmatrix} \alpha_0 \\ \underset{\sim}{\alpha}_3 \\ \underset{\sim}{\alpha}_4 \end{bmatrix} + \underset{\sim}{\varepsilon}. \qquad (3.4.3)$$

To get the estimates $\hat{\underset{\sim}{\alpha}}$ of $\underset{\sim}{\alpha}$ in (3.4.2) by the method of least squares, we have (with the usual assumptions on $\underset{\sim}{\varepsilon}$)

$$\hat{\underset{\sim}{\alpha}} = (\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'\underset{\sim}{y} . \qquad (3.4.4)$$

With the matrix $\underset{\sim}{W}$ defined as in (3.4.3), then

$$\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} N & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{\tilde{W}}'\underset{\sim}{\tilde{W}} & \underset{\sim}{\tilde{W}}'\underset{\sim}{C}_1 \\ \underset{\sim}{0} & \underset{\sim}{C}_1'\underset{\sim}{\tilde{W}} & \underset{\sim}{C}_1'\underset{\sim}{C}_1 \end{bmatrix} . \qquad (3.4.5)$$

Since the coefficients $z_{i.}^{(j)}$ in each group ($1 \le j \le q$) in each contrast ($1 \le i \le q-1$) are all equal in sign and magnitude and by the construction of the matrix $\underset{\sim}{\tilde{W}}$, that is, we shall construct the matrix $\underset{\sim}{\tilde{W}}$ so that each compartment of rows consists of a first-degree orthogonal design or

$$\sum_{u=N_{j-1}+1}^{N_j} w_{ui} = 0 \qquad (1 \le i \le p,\ 1 \le j \le q),$$

then

$$\underset{\sim}{\tilde{W}}'\underset{\sim}{C}_1 = \underset{\sim}{0}. \qquad (3.4.6)$$

Thus the matrix $\underset{\sim}{W}'\underset{\sim}{W}$ in (3.4.5) is a k x k diagonal matrix of diagonal matrices and the inversion of $\underset{\sim}{W}'\underset{\sim}{W}$ to obtain the estimates $\hat{\underset{\sim}{\alpha}}$ is a very simple operation.

To complete the operation in (3.4.4), we have the matrix

$$\underset{\sim}{W}'\underset{\sim}{y} = \begin{bmatrix} \sum_{u=1}^{N} y_u \\ \tilde{\underset{\sim}{W}}'\underset{\sim}{y} \\ \underset{\sim}{C}_1'\underset{\sim}{y} \end{bmatrix} \qquad (3.4.7)$$

where the matrix (vector) $\underset{\sim}{W}'\underset{\sim}{y}$ is a k x 1 matrix whose elements are linear functions of the observations. For example, the i^{th} element of $\underset{\sim}{C}_1'\underset{\sim}{y}$ is L_i. Therefore, if we substitute (3.4.5), (3.4.6) and (3.4.7) in (3.4.4), the vector of estimates is

$$\hat{\underset{\sim}{\alpha}} = \begin{bmatrix} \frac{1}{N}\sum_{u=1}^{N} y_u \\ (\tilde{\underset{\sim}{W}}'\tilde{\underset{\sim}{W}})^{-1}\tilde{\underset{\sim}{W}}'\underset{\sim}{y} \\ (\underset{\sim}{C}_1'\underset{\sim}{C}_1)^{-1}\underset{\sim}{C}_1'\underset{\sim}{y} \end{bmatrix} = \begin{bmatrix} \hat{\alpha}_0 \\ \hat{\underset{\sim}{\alpha}}_3 \\ \hat{\underset{\sim}{\alpha}}_4 \end{bmatrix} . \qquad (3.4.8)$$

In (3.4.8), $\hat{\underset{\sim}{\alpha}}_3$ corresponds to the estimates of the effects of the design variables and $\hat{\underset{\sim}{\alpha}}_4$ the estimates of the effects of the contrast terms. We can see that the estimates $\hat{\underset{\sim}{\alpha}}_3$ are unaffected by the addition of the q terms to the model since $\hat{\underset{\sim}{\alpha}}_3=(\tilde{\underset{\sim}{W}}'\tilde{\underset{\sim}{W}})^{-1}\tilde{\underset{\sim}{W}}'\underset{\sim}{y}$ is the same in both (4) of Section 3.2 and equation (3.4.8). Furthermore, since the parameters $\underset{\sim}{\beta}$ in the model of the intermediate variables are linear functions of the elements of $\underset{\sim}{\alpha}_3$, the estimates $\hat{\underset{\sim}{\beta}}$ are also unaffected by the contrast effects.

Let us next consider an example which should help to

clarify the material in Sections 3.3 and 3.4. At the conclusion of the example, we shall briefly discuss, before going into Chapter 4, what we have achieved by the addition of the contrast terms to the model.

An example.

In this example, we assume that in the manufacture of an aquatic paint used for naval vessels, the composition of the paint consists of components from three distinct categories. The components in category 1 might be the basic constituents used in the primary formulation of the paint. The components in category 2 could be solvents enabling the easy application of the paint onto rough surfaces and the components from category 3 might be assumed to contribute to the durability of the paint by providing water resistant characteristics.

Since the manufacturing process cannot be shut down for the running of the experiments, a pilot plant operation will be used. The experiments however will take approximately one full day (24 hours) since only one machine can be used, and therefore the execution of the experimental program will be exposed to three different shifts of operators. To eliminate any variation that might exist between the shifts of operators, we shall set the experimental design up in three parts using two contrasts to measure the variability between the shifts.

Let $\tilde{\underset{\sim}{W}}_i$ be the matrix of design variables to be run during the i^{th} shift. If $N_1=N_2=N_3=N/3$, then the contrast coefficients will be constructed as in (3.3.6), (3.3.6A) and (3.3.8) and the matrix $\underset{\sim}{W}$ containing the design variables will be of the form

$$\underset{\sim}{W} = \begin{bmatrix} & \tilde{\underset{\sim}{W}}_1 & \underset{\sim}{1}_{N/3} & \underset{\sim}{1}_{N/3} \\ \underset{\sim}{1}_N & \tilde{\underset{\sim}{W}}_2 & -\underset{\sim}{1}_{N/3} & \underset{\sim}{1}_{N/3} \\ & \tilde{\underset{\sim}{W}}_3 & \underset{\sim}{0}_{N/3} & -\underset{\sim}{2}_{N/3} \end{bmatrix}$$

where $\underset{\sim}{1}_N$ is an N x 1 column vector of 1's. The last two N x 1 column vectors in the matrix $\underset{\sim}{W}$ consist of the N/3 x 1 column vectors $\kappa \underset{\sim}{j}_{N/3}$ where $\kappa=0,\pm1,-2$. The construction of the matrices $\tilde{\underset{\sim}{W}}_i$ $(1\leq i\leq 3)$ will be discussed in Chapters 5 and 6.

The matrix $\underset{\sim}{W}$ shows only the plan of the experiment in the design variables. To obtain the experimental levels of the mixture components, we must use equation (3.2.15). The form of the matrix $\underset{\sim}{T}_1$ to be used in (3.2.15) will depend on n_1, n_2, n_3 the numbers of components in each of the three categories and can be constructed easily as shown in Appendices D and E.

By the inclusion of blocks or groups of observations using the theory of orthogonal contrasts, we have not at all influenced the estimation of the polynomial coefficients in the models of the design variables or intermediate variables. What has been achieved is that of limiting the experimental error to the error occuring only within the groups of observations. This is because the estimates of the parameters

associated with the contrast terms in the model give a measure of the between groups variation (see (3.3.9) and (3.3.10)), and thereby remove this variation from the error sum of squares in the analysis of variance. This is seen by fitting the models in (3.2.6) and (3.4.2). That is, if we fit the model (3.2.6), the breakdown of sum of squares in the analysis of variance table is

$$\begin{aligned} &\text{Sum of Squares of Regression} = \hat{\underset{\sim}{\alpha}}_3'\tilde{\underset{\sim}{W}}'\underset{\sim}{y} \\ &\text{Sum of Squares of Error} \quad\quad = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{\alpha}}_3'\tilde{\underset{\sim}{W}}'\underset{\sim}{y} \\ \hline &\text{Total Sum of Squares} \quad\quad = \underset{\sim}{y}'\underset{\sim}{y} \end{aligned} \quad .$$

Now if we fit the model (3.4.2), the analysis of variance table is

$$\begin{aligned} &\text{Sum of Squares of Regression*} = \hat{\underset{\sim}{\alpha}}'\underset{\sim}{W}'\underset{\sim}{y} \\ &\text{Sum of Squares of Error*} \quad\quad = \underset{\sim}{y}'\underset{\sim}{y} - \hat{\underset{\sim}{\alpha}}'\underset{\sim}{W}'\underset{\sim}{y} \\ \hline &\text{Total Sum of Squares} \quad\quad = \underset{\sim}{y}'\underset{\sim}{y} \end{aligned} \quad .$$

The formula for Sum of Squares of Regression* however is

$$\begin{aligned} \text{SSRegression*} = \hat{\underset{\sim}{\alpha}}'\underset{\sim}{W}'\underset{\sim}{y} &= [\hat{\alpha}_0, \hat{\underset{\sim}{\alpha}}_3', \hat{\underset{\sim}{\alpha}}_4'] \begin{bmatrix} \underset{\sim}{j}_N' \\ \tilde{\underset{\sim}{W}}' \\ \underset{\sim}{C}_1' \end{bmatrix} \underset{\sim}{y} \\ &= \hat{\alpha}_0 \sum_{u=1}^{N} y_u + \hat{\underset{\sim}{\alpha}}_3'\tilde{\underset{\sim}{W}}'\underset{\sim}{y} + \hat{\underset{\sim}{\alpha}}_4'\underset{\sim}{C}_1'\underset{\sim}{y} \\ &= \text{SSRegression} + \hat{\alpha}_0 \sum_{u=1}^{N} y_u + \hat{\underset{\sim}{\alpha}}_4'\underset{\sim}{C}_1'\underset{\sim}{y}. \end{aligned}$$

Since the sum of squares is non-negative, that is,

$$\hat{\alpha}_0 \sum_{u=1}^{N} y_u + \hat{\underset{\sim}{\alpha}}_4' \underset{\sim}{C}_1' \underset{\sim}{y} \geq 0,$$

then

Sum of Squares Regression* ≥ Sum of Squares Regression

and hence,

Sum of Squares of Error* ≤ Sum of Squares of Error. (3.4.14)

We have justified using compartments or groups of observations only so far as to help eliminate variation arising from different environmental (time, shifts of workers, etc.) conditions. Actually, in attempting to explore the response as an unknown function of several variables, an experimenter's strategy usually generates sequences of experiments that naturally fall into separate groups. For example, an experimenter's first step is usually to approximate the response using a first-degree model. Then by some measure of the lack of fit of the model, if a non-planar response is observed, the experimental design is augmented by one or more additional groups of experiments to permit the estimation of the coefficients in a second-degree model. Thus the experimental program progresses sequentially, using groups of experimental points. In fact, it will be shown in Chapter 7 that each compartment of design points used for second-degree designs is itself a first-degree design and the only difference in the design matrix is the addition of center point replicates.

IV. DESIGN CRITERIA AND AUGMENTED REGION OF INTEREST.

In the last chapter, the transformation from the system of the mixture components to a system involving design variables was shown. With the design variables then, experimental designs will be constructed using known methods. It was also mentioned at the beginning of Chapter 3 that when selecting criteria for the purpose of design optimization, the attention will be on the predictor $\hat{y}$. A criterion such as the average mean square error of $\hat{y}$, originally introduced by Box and Draper [3] for choosing appropriate response surface designs, is based on the properties (variance and bias) of $\hat{y}$ at each point in the factor space. Since this criterion can easily be applied to the polynomial models which will be used in this paper, we now consider this criterion.

4.1 The average mean square error of $\hat{y}$.

As in Chapter 3, let $\hat{y}(\underset{\sim}{w})$ represent the response estimated at the point $\underset{\sim}{w}$. The objective will be to choose the design matrix $\underset{\sim}{D}_w$ of N rows and k-q=p columns so that the difference $\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})$, where $\eta(\underset{\sim}{w})$ is the value of the true polynomial function at $\underset{\sim}{w}$, is small or minimized over the region of interest. By choosing the design matrix $\underset{\sim}{D}_w$ is meant specifying the levels of the p variables w_i to be run in the N experiments. The region of interest (unit sphere) will be denoted by R.

The measure of "closeness" which will be used at a

particular point $\underset{\sim}{w}$ is

$$E\,[\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})]^2. \tag{4.1.1}$$

Over the region of interest R, the average of (4.1.1) will be used, that is,

$$\Omega \int_R E[\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})]^2 \, d\underset{\sim}{w} \tag{4.1.2}$$

where

$$\Omega^{-1} = \int_R d\underset{\sim}{w}. \tag{4.1.3}$$

Now, in order to talk about designs (for possible comparisons) which do not contain the same number of points and also to insure that the criterion is independent of the variance σ^2 of the observations, the criterion (4.1.2) is placed on a "per observation" basis and denoted by

$$J = \frac{N\Omega}{\sigma^2} \int_R E[\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})]^2 \, d\underset{\sim}{w}. \tag{4.1.4}$$

The difference $\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})$ can be partitioned into

$$\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w}) = \{\hat{y}(\underset{\sim}{w}) - E\hat{y}(\underset{\sim}{w})\} + \{E\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})\}, \tag{4.1.5}$$

which enables us to separate J into two parts, that is,

$$J = \frac{N\Omega}{\sigma^2} \int_R \text{Var}\, \hat{y}(\underset{\sim}{w}) \, d\underset{\sim}{w} + \frac{N\Omega}{\sigma^2} \int_R \{E\hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})\}^2 \, d\underset{\sim}{w}$$

$$= V + B. \tag{4.1.6}$$

In this last expression (4.1.6) for J, the quantities V and B are called the average variance and the average squared bias respectively, where average means averaged over the region of interest R.

In the material that follows in this section, we shall

assume that q=2 and the approximating function or fitted model is a polynomial of degree 1 in the $\underset{\sim}{w}$, i.e.,

$$y(\underset{\sim}{w}) = \underset{\sim}{w}_1' \underset{\sim}{\alpha}_1 + \epsilon \qquad (4.1.7)$$

where $\underset{\sim}{w}_1'=(1, w_1, \ldots, w_p, \pm 1)$, $\underset{\sim}{\alpha}_1'=(\alpha_0, \alpha_1, \ldots, \alpha_p, \alpha_{L_1})$ and α_{L_1} is the parameter associated with the contrast term. It will also be assumed that the true polynomial function is of the form

$$\eta(\underset{\sim}{w}) = \underset{\sim}{w}_1' \underset{\sim}{\alpha}_1 + \underset{\sim}{w}_2' \underset{\sim}{\alpha}_2 \qquad (4.1.8)$$

where the vector $\underset{\sim}{w}_2$ contains terms of degree 2. The reason for using models of only degree 1 and 2 above is that we are assuming it is unlikely that a model of third-degree or higher is needed to adequately describe the mixture system especially in our reduced region of interest. Even in the extended region of interest (Section 4.3), the fitted model will be at most second degree.

Now, to minimize the expression (4.1.6), we separately consider the following three conditions:

(i) the assumed degree of the fitted model is adequate but error arises owing to the variance of the predictor $\hat{y}$,

(ii) the variance of $\hat{y}$ is assumed to be negligible and therefore if error is present it is due only to a misclassification of the model,

(iii) neither the variance nor the bias error can be ignored and therefore we want to minimize the quantity J = V + B.

Before we discuss the minimization of J corresponding to the above three conditions, we define what are meant by

the moments of the design and region moments (see for example, Box and Draper [3]). This discussion will be useful since both types of moments will be used extensively in the development of the theory of design construction for the minimization of the average mean square error of $\hat{y}$.

4.2 Design moments and region moments.

In order to specify the distribution of the design points in the experimental region, we use the moments of the design. In Section 3.4, the $\underset{\sim}{W}'\underset{\sim}{W}$ matrix was used for obtaining the estimates of the parameters in the general linear model. Now the quantities in the matrix $\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W}$ (the moment matrix), where N is the number of experiments, are the moments of the design up to the second degree. The usual notation in response surface work is to designate the moments in square brackets as

$$[i] = \frac{1}{N}\sum_{u=1}^{N} w_{ui} \quad , \quad [ii] = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^{2} \quad , \quad [ij] = \frac{1}{N}\sum_{u=1}^{N} w_{ui}w_{uj} . \qquad (4\cdot2\cdot1)$$

Note that if the model contains terms up to second degree, the moment matrix will contain moments up to fourth degree such as [iiii], [iijj], etc.

By the particular method with which we have defined the elements of the design matrix, that is, $w_{1i}+w_{2i}+\ldots+w_{Ni}=0$ $(1\leq i\leq p)$, then

$$[ii] = \frac{1}{N} \sum_{u=1}^{N} w_{ui}^2 \qquad (4.2.2)$$

is a measure of the spread or variability of the design points associated with the i^{th} variable. In addition, had we specified that $[ii] = 1$ as in [3, 6], then

$$\frac{[iii]}{[ii]^{3/2}} = \frac{1}{N} \sum_{u=1}^{N} w_{ui}^3 \qquad (4.2.3)$$

is a measure of skewness, i.e., a measure of the extent to which the marginal distribution of the pattern of design points for the i^{th} variable is skew or symmetric. Also, if $[ii] = 1$, then

$$\frac{[iiii]}{[ii]^2} = \frac{1}{N} \sum_{u=1}^{N} w_{ui}^4 \qquad (4.2.4)$$

is a measure of kurtosis, i.e., a measure of the extent to which there is a tendency to a uniform distribution of points or to a concentration of points at the center and at the extremes.

Since the w_{ui} ($1 \leq u \leq N$, $1 \leq i \leq p$) are chosen by the experimenter, the experimenter has at his freedom the choice of the moments of the particular design. However, once the form of the moment matrix is specified, the distribution of the design points is determined.

Region moments.

In Section 4.3, the radius of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the convex polytope is

derived. Denote this radius by ρ^*. Then with either the unit sphere or the largest sphere defined as our region of interest, we assume that the "center of gravity" of the design points is at the center of the region defined by

$$\sum_{i=1}^{p} w_{ui}^2 \leq \rho^2 \qquad (4\cdot2\cdot5)$$

where $\rho=1$ or $\rho=\rho^*\geq 1$. Let $\rho=\rho^*$. If $\rho^* > 1$, we lift the restriction that the design points must fall within or on the unit sphere. On the other hand, we shall limit the position of the design points to either within or on the sphere of radius ρ^* and denote this augmented sphere by R^*.

For the region defined by (4·2·5), Dirichlet multiple integrals can be used. That is to say,

$$\int_{R^*} w_1^{\delta_1} w_2^{\delta_2} \ldots w_p^{\delta_p} d\underset{\sim}{w} = \frac{\Gamma\left(\frac{\delta_1+1}{2}\right)\Gamma\left(\frac{\delta_2+1}{2}\right)\cdots\Gamma\left(\frac{\delta_p+1}{2}\right)}{\Gamma\left\{\sum_{i=1}^{p}\frac{(\delta_i+1)}{2}+1\right\}}\,\rho^{*\sum_{i=1}^{p}(\delta_i+1)} \qquad (4.2.6)$$

unless any δ_i is odd in which case the value of the integral is zero. Now Box and Draper [3] define the matrix of region moments as

$$\underset{\sim}{\mu}_{ef} = \int_R \underset{\sim}{w}_e \underset{\sim}{w}_f' \, d\underset{\sim}{w} \qquad (4.2.7)$$

where $\underset{\sim}{w}_e$, $\underset{\sim}{w}_f$ are the vectors of terms in the models of degree e, f (e,f = 1,2), R is the region of interest and Ω is defined by (4.1.3). To find Ω^{-1} in the formula (4.1.3)

in Section 4.1, we have using (4.2.6),

$$\Omega^{-1} = \int_{R^*} w_1^0 \, w_2^0 \, \ldots \, w_p^0 \, dw$$

$$= \frac{\left[\Gamma\left(\frac{1}{2}\right)\right]^p}{\Gamma\left(\frac{p}{2}+1\right)} \rho^{*p} \, . \qquad (4.2.8)$$

In a similar manner, from (4.2.7),

$$\mu_{11} = \Omega \int_{R^*} \begin{bmatrix} 1 & w_1 & \ldots & w_p & \pm 1 \\ w_1 & w_1^2 & \ldots & w_1 w_p & \pm w_1 \\ \vdots & \vdots & & \vdots & \vdots \\ w_p & w_p w_1 & \ldots & w_p^2 & \pm w_p \\ \pm 1 & \pm w_1 & \ldots & \pm w_p & 1 \end{bmatrix} dw \qquad (4.2.9)$$

$$= \begin{bmatrix} 1 & 0' & \pm 1 \\ 0 & \frac{\rho^{*2}}{p+2} I_p & 0 \\ \pm 1 & 0' & 1 \end{bmatrix} \qquad (4.2.10)$$

since $w_1' = (1, w_1, \ldots, w_p, \pm 1)$ and where I_p is a p^{th} order identity matrix. The signs in the last row and last column (with the exception of 1) of the matrix in (4.2.9) would all be plus or minus depending on whether the sign of the last element in w_1' is plus or minus. In addition, the matrices μ_{21} and μ_{22} are

$$\mu_{12}' = \mu_{21} = \Omega \int_{R^*} w_2 w_1' \, dw$$

$$\underset{\sim}{\mu}_{21} = \frac{\rho^{*2}}{p+2} \begin{bmatrix} \underset{\sim}{j}_p & \underset{(p\times p)}{\underset{\sim}{0}} & \pm\underset{\sim}{j}_p \\ \\ \underset{(\frac{p}{2})\times 1}{\underset{\sim}{0}} & \underset{\sim}{0} & \underset{\sim}{0} \end{bmatrix} \tag{4.2.11}$$

and

$$\underset{\sim}{\mu}_{22} = \Omega \int_{R^*} \underset{\sim}{w}_2 \underset{\sim}{w}_2' \, d\underset{\sim}{w}$$

$$= \frac{\rho^{*4}}{(p+2)(p+4)} \begin{bmatrix} 2\underset{\sim}{I}_p + \underset{\sim}{j}_p \underset{\sim}{j}_p' & \underset{p\times(\frac{p}{2})}{\underset{\sim}{0}} \\ \\ \underset{(\frac{p}{2})\times p}{\underset{\sim}{0}} & \underset{\sim}{I}_{(\frac{p}{2})} \end{bmatrix} \tag{4.2.12}$$

where

$$\underset{\sim}{w}_2' = (w_1^2, \ldots, w_p^2, w_1 w_2, \ldots, w_{p-1} w_p)$$

and $\underset{\sim}{j}_p$ is a p x 1 column vector of ones. If

$$\underset{\sim}{w}_2' = (w_1^2, \ldots, w_p^2, \sqrt{2}w_1 w_2, \ldots, \sqrt{2}w_{p-1} w_p)$$

as in Chapter 7, then (4.2.12) is the same except that $\underset{\sim}{I}_{(\frac{p}{2})}$ is replaced with $2\underset{\sim}{I}_{(\frac{p}{2})}$.

We now consider the three cases of minimizing J mentioned in Section 4.1. The first case is where B is assumed to be zero and therefore the emphasis is on minimizing the average variance only. The second case is that of minimizing

J = B, i.e., the variance of $\hat{y}$ is assumed to be negligible. The third case which will be covered in Appendix G is where we minimize the variance plus the square of the bias. In [3], designs constructed for the purpose of minimizing J in the first two cases are called "all-variance" and "all-bias" designs respectively.

Case (i) Minimizing average variance when the bias is assumed to be negligible.

One of the objectives of most planned experiments is to minimize the variance of the predicted quantity for future values of the concomitant variables when the bias can be ignored. The "all-variance" design minimizes the variance criterion

$$V = \frac{N\Omega}{\sigma^2} \int_R \text{Var } \hat{y}(\underset{\sim}{w})\ d\underset{\sim}{w} \qquad (4.2.13)$$

Since the fitted model in (4.2.13) is of degree 1, the model is written as

$$\hat{y}(\underset{\sim}{w}) = \underset{\sim}{w}_1' \ \hat{\underset{\sim}{\alpha}}_1.$$

From Section 3.5, $\hat{\underset{\sim}{\alpha}}_1 = (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{y}$, and therefore the estimated value can be written as

$$\hat{y}(\underset{\sim}{w}) = \underset{\sim}{w}_1' \ (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{y}\ , \qquad (4.2.14)$$

where $\underset{\sim}{W}_1$ corresponds to the Nxk matrix of independent variables

associated with the first-degree model. With the usual assumptions on $\underset{\sim}{\varepsilon}$, i.e., $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$, $E(\underset{\sim}{\varepsilon}\underset{\sim}{\varepsilon}') = \sigma^2 \underset{\sim}{I}_N$, the variance of the predictor $\hat{y}(\underset{\sim}{w})$ is

$$\text{Var } \hat{y}(\underset{\sim}{w}) = \sigma^2 \, \underset{\sim}{w}_1' (\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1} \underset{\sim}{w}_1 \, . \qquad (4.2.15)$$

Now if we put (4.4.15) into the formula (4.2.13) for V, we have

$$V = N\Omega \int_R \underset{\sim}{w}_1' \, (\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1} \underset{\sim}{w}_1 \, d\underset{\sim}{w}$$

$$= N\Omega \int_R \text{trace} \, [\, (\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1} \underset{\sim}{w}_1 \underset{\sim}{w}_1'] d\underset{\sim}{w}. \qquad (4.2.16)$$

However, since $\underset{\sim}{W}_1' \underset{\sim}{W}_1$ is constant,

$$V = \text{trace} \, [\, N(\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1} \, \{ \, \Omega \int_R \underset{\sim}{w}_1 \underset{\sim}{w}_1' \, d\underset{\sim}{w} \, \} \,]$$

$$= \text{trace} \, \{ \, N(\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1} \underset{\sim}{\mu}_{11} \, \} \, . \qquad (4.2.17)$$

In (4.2.17), the elements of $(\underset{\sim}{W}_1' \underset{\sim}{W}_1)^{-1}$ are the variances of the estimates $\hat{\alpha}_i$ of the $\alpha_i (i=1,2,\ldots,p,0,L_1)$ in the model $\underset{\sim}{y} = \underset{\sim}{w}_1' \underset{\sim}{\alpha}_1 + \underset{\sim}{\varepsilon}$ and if we note the construction of $\underset{\sim}{\mu}_{11}$ from (4.2.10), then (4.2.17) can be rewritten as

$$V = N \, [\, \text{Var } \hat{\alpha}_0 + \frac{\rho^{*2}}{p+2} \sum_{i=1}^{p} \text{Var } \hat{\alpha}_i + \text{Var } \hat{\alpha}_{L_1} \,] \qquad (4.2.18)$$

where $\hat{\alpha}_{L_1}$ is the estimate of the constrast parameter (we have been assuming q=2), and R* is the region of interest. Box [1]

showed that the first p+1 terms in (4.2.18) are minimized when the design in $\underset{\sim}{W}_1$ is orthogonal. That is, minimum variance first-degree designs are provided by any set of orthogonal vectors for which $\sum_{u=1}^{N} w_{ui} = 0$ for $i = 1,2,\ldots,p$.

Case (ii) Minimizing average squared bias when the variance is assumed to be negligible.

In Appendix 1, [3], Box and Draper show that for bias alone to be minimized, a sufficient condition is that the design moments must be equal to the region moments up to and including degree (d_1+d_2), where d_1 is the degree of the fitted model and d_2 is the degree of the true polynomial function and $d_2 \geq d_1$. In our case, $d_1=1$ and $d_2=2$, i.e.,

$$\hat{y}(\underset{\sim}{w}) = \underset{\sim}{w}_1' \, \hat{\underset{\sim}{\alpha}}_1$$

and

$$\eta(\underset{\sim}{w}) = \underset{\sim}{w}_1' \, \underset{\sim}{\alpha}_1 + \underset{\sim}{w}_2' \, \underset{\sim}{\alpha}_2.$$

When this is true,

$$E(\hat{\underset{\sim}{\alpha}}_1) = \underset{\sim}{\alpha}_1 + (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1} \underset{\sim}{W}_1'\underset{\sim}{W}_2 \, \underset{\sim}{\alpha}_2$$

$$= \underset{\sim}{\alpha}_1 + \underset{\sim}{A} \, \underset{\sim}{\alpha}_2 \, ,$$

where $\underset{\sim}{A} = (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1} \underset{\sim}{W}_1'\underset{\sim}{W}_2$ is called the alias matrix.

In Section 4.1, equation (4.1.6), the average squared bias is expressed as

$$B = \frac{N\Omega}{\sigma^2} \int_R [E\ \hat{y}(\underset{\sim}{w}) - \eta(\underset{\sim}{w})]^2\ d\underset{\sim}{w}.$$

If we substitute $\hat{y}(\underset{\sim}{w})$, $\eta(\underset{\sim}{w})$ and $\underset{\sim}{A}$ into the formula for B, then

$$B = \frac{N\Omega}{\sigma^2} \int_R \underset{\sim}{\alpha}_2'[\underset{\sim}{w}_1'\underset{\sim}{A} - \underset{\sim}{w}_2'][\underset{\sim}{w}_1'\underset{\sim}{A} - \underset{\sim}{w}_2']\underset{\sim}{\alpha}_2\ d\underset{\sim}{w}$$

$$= \frac{N}{\sigma^2}\ \underset{\sim}{\alpha}_2'\ \underset{\sim}{\Delta}\ \underset{\sim}{\alpha}_2 \qquad (4.2.19)$$

where

$$\underset{\sim}{\Delta} = \underset{\sim}{A}'\underset{\sim}{\mu}_{11}\underset{\sim}{A} - \underset{\sim}{\mu}_{12}'\underset{\sim}{A} - \underset{\sim}{A}'\underset{\sim}{\mu}_{12} + \underset{\sim}{\mu}_{22}\ .$$

Now if we partition $\underset{\sim}{\Delta}$, we have

$$B = \frac{N}{\sigma^2}\ \underset{\sim}{\alpha}_2'\ [(\underset{\sim}{\mu}_{22} - \underset{\sim}{\mu}_{12}\underset{\sim}{\mu}_{11}^{-1}\underset{\sim}{\mu}_{12})$$

$$+ (\underset{\sim}{A} - \underset{\sim}{\mu}_{11}^{-1}\underset{\sim}{\mu}_{12})'\underset{\sim}{\mu}_{11}(\underset{\sim}{A} - \underset{\sim}{\mu}_{11}^{-1}\underset{\sim}{\mu}_{12})]\ \underset{\sim}{\alpha}_2. \qquad (4.2.20)$$

If we set

$$\underset{\sim}{A} = \underset{\sim}{\mu}_{11}^{-1}\underset{\sim}{\mu}_{12}\ ,$$

the second term in (4.2.20) is eliminated. Box and Draper state that in particular, B is minimized when $\underset{\sim}{W}_1'\underset{\sim}{W}_1 = \underset{\sim}{\mu}_{11}$ and $\underset{\sim}{W}_1'\underset{\sim}{W}_2 = \underset{\sim}{\mu}_{12}$ which is just a statement that the moments of the design equal the moments of the region R (or R*) up to and including degree (d_1+d_2).

Case (iii) Minimizing the average mean square error: J=V+B.

Here the emphasis is on a design which will guard against not only model misclassification but excessive variance of $\hat{y}$. The minimization of J=V+B however will not be discussed now since this subject is covered at some length in Appendix G.

4.3 Extended region of interest: Largest sphere centered at the point of interest that will fit inside the factor space.

Suppose the region in the $\underline{w}$ space in which the experiments can be performed is called the operability region. The operability region is the (k-q)-dimensional design space bounded by the extremities of the convex polytope. In some cases, the experimenter may wish to explore the entire operability region, but we have assumed in the present paper that this is not the case. We have stated that a particular group of experiments is performed in order to explore a region of interest R entirely contained within the region of operation. The reason for this as we said before is that situations frequently arise where the interest is attached only to some reduced region and therefore certain parts of the region of operation can be ignored.

Since it is impossible to forsee all conceivable choices for the shape of the region of interest R, we have restricted the development in this paper to the case where R

is spherical. The reasons for this are two-fold. In the first place, this is necessitated by the complexity of the operability region or shape of the convex polytope as the number of categories as well as the number of components increases. We saw in the last section that the region moments for a spherical region can easily be calculated for various values of q and k using Dirichlet multiple integrals. Secondly, by the scaling convention used in this paper, that is, the specification of an ellipsoidal region of interest for the mixture components and then the setting up of the spherical region of interest in the design variables. For instance, if we refer back to Section 3.1, the form (3.1.2) describes the intermediate variables v_j $(1 \leq j \leq k)$ as linear functions of the mixture components. Then by specifying the v_j $(1 \leq j \leq k)$ as falling within or on the sphere $v_1^2 + \ldots + v_k^2 \leq 1$, we not only defined the region of interest as being a unit sphere in the metric of the v_j but by using an orthogonal transformation matrix $\underset{\sim}{T}$, the region of interest is a unit sphere also in the metric of the w_i $(1 \leq i \leq k-q)$, the design variables.

Since the region of interest may at times be located (depending on $\underset{\sim}{x}_0$ and $\underset{\sim}{H}$) near a boundary of the region of operation, it might be well to determine how much flexibility we have before reaching the boundary. This flexibility can be measured by calculating the radius of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the factor space or

convex polytope and comparing it with $\rho=1$, the radius of the unit sphere. Then if the need arises, we may be able to extend our interest outside the previously defined region of interest. Arising needs could be the result of an underestimate of the interval of interest for one or more of the mixture components or simply the construction of a design of the second degree (Chapter 7) which consists of concentric sets of points and it is desired to construct a set of points which has radius greater than unity.

Let us denote the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the convex polytope by R* as in Section 4.2. In order to determine the length of the radius of the (k-q)-dimensional sphere, we have to find the distance from the center $\underset{\sim}{w}=\underset{\sim}{0}$ to the closest boundary of the polytope. The case where q=2 will be shown now and an extension to q=3 will be shown in Chapter 6.

The dimensionality of the boundary (extremity) closest to the point $\underset{\sim}{w}=\underset{\sim}{0}$ will be 1 less than the dimensionality of the convex polytope. This is proved in Appendix C. For example, in two dimensions, the distance from the point $\underset{\sim}{w}=\underset{\sim}{0}$ to the closest edge (dimensionality 1) is less than the distance to the closest vertex (dimensionality 0). In three dimensions, the distance from the point $\underset{\sim}{w}=\underset{\sim}{0}$ to the closest face (dimensionality 2) is less than the distance to the closest edge or closest vertex.

This proposition would be obvious if "less" were

replaced by "not more" and in fact this is all we require for our purposes, although the stronger result is of geometrical interest in its own right.

In the case of two categories where n is the number of components in category 1 and k is the total number of components, an extremity of dimensionality k-3 was defined in Section 2.3 by either

$$x_i = 0, \quad \sum_{\substack{i'=1 \\ i' \neq i}}^{n} x_{i'} = \frac{1}{2}, \quad \sum_{j=n+1}^{k} x_j = \frac{1}{2} \quad (i=1,\ldots,n), \quad (4.3.1)$$

or

$$\sum_{i=1}^{n} x_i = \frac{1}{2}, \quad x_j = 0, \quad \sum_{\substack{j'=n+1 \\ j' \neq j}}^{k} x_{j'} = 0 \quad (j=n+1,\ldots,k), \quad (4.3.2)$$

Let the closest extremity be defined by (4.3.2) where j=n+1. Since the v_i's have the same metric as the w_i's, we can work with the v_i's. From (3.1.2), we know that

$$v_i = \frac{x_i - x_{0i}}{h_i} \quad \text{for } 1 \leq i \leq k,$$

and therefore in the v_i, the restrictions corresponding to (4.3.2) are

$$\sum_{i=1}^{n} (h_i v_i + x_{0i}) = \frac{1}{2}, \quad (4.3.3)$$

$$h_{n+1} v_{n+1} + x_{0n+1} = 0, \quad (4.3.4)$$

and

$$\sum_{j'=n+2}^{k} (h_{j'}v_{j'} + x_{0j'}) = \frac{1}{2} . \qquad (4.3.5)$$

Now, every equation of the first degree

$$\sum_{i=1}^{\ell} a_i z_i = d$$

where z_i $(1 \le i \le \ell)$ is the i^{th} component of an ℓ-vector, represents either a point, line, plane, ..., (n-1)-flat of $n-1 \le \ell-1$ dimensions in space and every point, line, etc. can be so represented. If the coefficients a_i $(1 \le i \le \ell)$ are divided by the quantity

$$\left(\sum_{i=1}^{\ell} a_i^2 \right)^{\frac{1}{2}} ,$$

then the coefficients

$$b_i = \frac{a_i}{\left(\sum_{i=1}^{\ell} a_i^2 \right)^{\frac{1}{2}}} \qquad (1 \le i \le \ell)$$

are called the direction cosines of the plane, etc., and the equation can be written in the normal form as

$$\sum_{i=1}^{\ell} b_i z_i = d' \qquad (4.3.6)$$

where

$$d' = \frac{d}{\left(\sum_{i=1}^{\ell} a_i^2 \right)^{\frac{1}{2}}} . \qquad (4.3.7)$$

In addition, the square of the perpendicular distance from the origin to the plane defined by (4.3.6) is

$$d'^2 = \frac{d^2}{\sum_{i=1}^{\ell} a_i^2} . \qquad (4.3.7A)$$

Now if we put (4.3.3) in the normal form, we have

$$\sum_{i=1}^{n} \frac{h_i v_i}{\left(\sum_{i=1}^{n} h_i^2\right)^{\frac{1}{2}}} = \frac{\frac{1}{2} - \sum_{i=1}^{n} x_{0i}}{\left(\sum_{i=1}^{n} h_i^2\right)^{\frac{1}{2}}}$$

$$= 0 \qquad (4.3.8)$$

since $x_{01}+x_{02}+\ldots+x_{0n}= \frac{1}{2}$. From (4.3.8), we see upon substituting (4.3.8) in (4.3.7A), that the origin $\underset{\sim}{y}=\underset{\sim}{0}$ is a point on the boundary defined by (4.3.3). Putting (4.3.4) and (4.3.5) in the normal form respectively, we have

$$\frac{h_{n+1} v_{n+1}}{h_{n+1}} = \frac{x_{0n+1}}{h_{n+1}} , \qquad (4.3.9)$$

$$\sum_{j'=n+2}^{k} \frac{h_{j'} v_{j'}}{\left(\sum_{j'=n+2}^{k} h_{j'}^2\right)^{\frac{1}{2}}} = \frac{x_{0n+1}}{\left(\sum_{j'=n+2}^{k} h_{j'}^2\right)^{\frac{1}{2}}} . \qquad (4.3.10)$$

Then from (4.3.7A), the square of the distance from $\underset{\sim}{y}=\underset{\sim}{0}$ to the extremities defined by (4.3.9) and (4.3.10) respectively are

$$\frac{x_{0n+1}^2}{h_{n+1}^2} \tag{4.3.11}$$

and

$$\frac{x_{0n+1}^2}{\sum_{j'=n+2}^{k} h_{j'}^2} \quad . \tag{4.3.12}$$

Since we want the distance from $\underset{\sim}{y}=\underset{\sim}{0}$ to the boundary defined by (4.3.2) where the extremities defined by (4.3.9) and (4.3.10) are orthogonal (product of the direction cosines is zero) to one another, then the distance from $\underset{\sim}{y}=\underset{\sim}{0}$ to the boundary (4.3.2) is the distance to the intersection of (4.3.9) and (4.3.10). Using Pythagoras's theorem for (4.3.11) and (4.3.12), we have for the square of this distance

$$\rho^2 = x_{0n+1}^2 \left\{ \frac{1}{h_{n+1}^2} + \frac{1}{\sum_{j'=n+2}^{k} h_{j'}^2} \right\},$$

and therefore, the distance from $\underset{\sim}{y}=\underset{\sim}{0}$ to the extremity defined by (4.3.2) where j=n+1, is

$$\rho = x_{0n+1} \left\{ \frac{1}{h_{n+1}^2} + \frac{1}{\sum_{j'=n+2}^{k} h_{j'}^2} \right\}^{\frac{1}{2}} . \tag{4.3.13}$$

Since the extremities of dimensionality k-3 were defined by (4.3.1) and (4.3.2), then from (4.3.13) if we let

$$\rho_1 = \min_{1\le i\le n} x_{0i} \left\{ \frac{1}{h_i^2} + \frac{1}{\sum_{\substack{i'=1 \\ i'\neq i}}^{n} h_{i'}^2} \right\}^{\frac{1}{2}}$$

$$\rho_2 = \min_{n+1\le j\le k} x_{0j} \left\{ \frac{1}{h_j^2} + \frac{1}{\sum_{\substack{j'=n+1 \\ j'\neq j}}^{k} h_{j'}^2} \right\}^{\frac{1}{2}},$$

the radius of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ ($\underset{\sim}{v}=\underset{\sim}{0}$) that will fit inside the convex polytope is

$$\rho^* = \min \{\rho_1, \rho_2\}. \qquad (4.3.14)$$

With the use of (4.3.14) for the radius of the largest sphere, the experimenter now has a measure in terms of distance from the center of the design space upon which to place design points. In other words, the radius of the designs (the distance the design points lie from the center $\underset{\sim}{w}=\underset{\sim}{0}$ of the design) can now be expressed as some fraction of the radius of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the factor space. Therefore, in the next chapter concerning first-degree designs, we shall refer to this largest sphere (denoted by R*) as the extended region of interest. The reason for this as we pointed out at the beginning of this section is that the radius ρ^* is an upper bound for the distance that the design points may lie from the center of the designs (due to their symmetry with $\underset{\sim}{w}=\underset{\sim}{0}$) that we shall be using. Also, in using the largest spherical

region, we can always reduce this region to the original region of interest defined by (3.1.1) by putting $\rho^*=1$.

It should be mentioned however, that if the calculated ρ^* is much larger than unity, care should be taken in making the assumption of equivalent models (i.e., fitted models) in R^* and in the unit spherical region. If ρ^* is much larger than unity, it may be necessary to talk about the fitted models in the two regions as being of different degrees, i.e., fit a higher degree (second-degree) model in R^*.

V. FIRST-DEGREE DESIGNS.

In [3], Box and Draper show, when fitting a first-degree model when the true surface over the operability region is quadratic, that the average mean square error of $\hat{y}$ is minimized if one uses a first-degree orthogonal design which has third-degree moments equal to zero. In Chapter 3, in the reparametrization of the model $\underset{\sim}{y} = \underset{\sim}{V}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$, we stated that <u>any</u> k x (k-q) matrix $\underset{\sim}{T}_1$ could be used such that the k x k matrix $\underset{\sim}{T}$ is orthogonal. We now show that a class of designs exists which enable us not only to use any k x (k-q) matrix $\underset{\sim}{T}_1$ in the reparametrization but also the class of designs satisfies the first condition, i.e., the designs are orthogonal.

To describe another reparametrization of the model $\underset{\sim}{y} = \underset{\sim}{V}\,\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$, let $\underset{\sim}{T}^* = \underset{\sim}{T}\underset{\sim}{M}$ where $\underset{\sim}{M}$ is a k x k orthogonal (not an identity) matrix. Partition the matrices $\underset{\sim}{T}^*$, $\underset{\sim}{T}$ and $\underset{\sim}{M}$ in the following manner:

$$\underset{\sim}{T} = [\underset{\sim}{T}_1\ \underset{\sim}{T}_2],\quad \underset{\sim}{T}^* = [\underset{\sim}{T}_1^*\ \underset{\sim}{T}_2],\quad \underset{\sim}{M} = \begin{bmatrix} \underset{\sim}{M}_1 & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{I}_q \end{bmatrix}$$

where $\underset{\sim}{M}_1$ is a (k-q) x (k-q) orthogonal (not identity) matrix. Then

$$\underset{\sim}{T}^* = [\underset{\sim}{T}_1^*\ \underset{\sim}{T}_2] = [\underset{\sim}{T}_1\underset{\sim}{M}_1\ \underset{\sim}{T}_2]$$

or

$$\underset{\sim}{T}_1^* = \underset{\sim}{T}_1\underset{\sim}{M}_1.$$

If the resulting model using $\underset{\sim}{T}^*$ in the reparametrization is

$$\underset{\sim}{y} = \underset{\sim}{Z}\underset{\sim}{\delta}_3 + \underset{\sim}{\varepsilon},$$

where the matrix $\underset{\sim}{Z}$ is N x (k-q) of rank k-q and $\underset{\sim}{\delta}_3$ is defined in the same way as $\underset{\sim}{\alpha}_3$ in Section 3.2, then

$$\underset{\sim}{\delta}_3 = \underset{\sim}{T}_1^{*\prime}\underset{\sim}{\beta} = \underset{\sim}{M}_1'\underset{\sim}{T}_1'\underset{\sim}{\beta} = \underset{\sim}{M}_1'\underset{\sim}{\alpha}_3.$$

The models are

$$\begin{aligned}\underset{\sim}{y} &= \underset{\sim}{Z}\underset{\sim}{M}_1'\underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon} \\ &= \underset{\sim}{\tilde{W}}\underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}\end{aligned}$$

which shows that the elements of the matrix $\underset{\sim}{Z}$ are the result of some arbitrary rotation, defined by the matrix $\underset{\sim}{M}_1$, of the elements of $\underset{\sim}{\tilde{W}}$, i.e., $\underset{\sim}{Z} = \underset{\sim}{\tilde{W}}\underset{\sim}{M}_1$.

By stating then that any matrix $\underset{\sim}{T}_1$, such that $\underset{\sim}{T}$ is orthogonal, can be used, we shall need to use a class of designs for which the properties of $\hat{y}$ are invariant to an arbitrary rotation of the design points. One such class of designs satisfying this property is the class of rotatable designs and for first-degree, orthogonality and rotatability are synonomous.

Therefore, in this chapter (and Chapter 6), we restrict the choice of designs to first-degree rotatable designs in which the third-degree moments are zero. The latter property, that is, the third-degree moments of the design being zero, is a result of requiring the bottom half of the design matrix to be the negative replicate of the top half of the design matrix. That is to say, the levels of

the design variables in the bottom half are the reflections (in some cases), with respect to the center of gravity of the design, of the levels of the design variables in the top half of the design matrix. Box [1] points out that this insures that the third-degree moments of the design vanish. (We shall also verify this in Section 5.2.)

The reader familiar with the arrangements of design points associated with first and second-degree rotatable designs will not find the material in this chapter and Chapter 7 entirely new since we shall be working with simplex designs and two-level factorial designs. However, to the reader not familiar with the concept of rotatability of designs, this concept and the requirements for achieving rotatability of designs are discussed in Appendix D.

5.1 Description of the designs.

In setting-up a pattern of experimental design points, it must be understood that certain patterns of points might fail not only to provide accurate estimates of the parameters in the model but might not even allow certain parameters to be separately estimated at all. However, when we speak of a first-degree design, we shall mean an arrangement of design points which permit the separate estimation of all the parameters in a first-degree linear model.

In attempting to obtain a best distribution of experimental points, we shall keep the following three considerations in mind. The arrangement must be such that

(i) the variances of the estimated coefficients are as small as possible,

(ii) the biases in the estimated coefficients which might occur if the first-degree model is representationally inadequate should be as small as possible,

(iii) since these designs are the first step in any type of sequential experimentation, we must concentrate our efforts only on first-degree designs which can easily be augmented by the addition of points to second-degree designs.

When q=2, the first-degree model in the design variables is of the form

$$\underset{\sim}{y} = \underset{\sim}{W}\,\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} \qquad (5.1.1)$$

where

$$\underset{\sim}{W} = \begin{bmatrix} \underset{\sim}{j}_N & \underset{\sim}{D}_w & \underset{\sim}{j}_{N/2} \\ & & -\underset{\sim}{j}_{N/2} \end{bmatrix}, \qquad (5.1.2)$$

and $\underset{\sim}{j}_{n'}$, (n'=N, N/2) is an n' x 1 column vector of ones. In (5.1.2), $\underset{\sim}{D}_w$ is the design matrix of the w_i's ($1 \le i \le k-2$) and the second-degree moments of the design are defined by $\lambda_2 = \frac{1}{N}(w_{1i}^2 + \ldots + w_{Ni}^2)$ for $1 \le i \le k-2$. Note that the second-degree moments of a first-degree rotatable design are all equal to the same value. This value is denoted by λ_2.

In the material that follows, two types of designs will be considered and for both designs, it will be shown that the third-degree moments vanish. Let p = k-2. The first type of design is the <u>double-simplex</u> (to be defined

shortly) where there are N=2(p+1) experimental points and the design points lie at the vertices of two regular p-dimensional simplices inscribed in the region of interest R*. The second type of design considered is the scaled two-level factorial where $N=2 \times 2^p$ or some one-half fraction thereof. The design matrix of the scaled two-level factorial is of the form

$$\underset{\sim}{D}_w = c\begin{bmatrix} \underset{\sim}{F} \\ -\underset{\sim}{F} \end{bmatrix}, \qquad (5.1.3)$$

where c is a scale factor such that $\lambda_2 = c^2$ is not too large for the design points to fit in the restricted factor space. The $\frac{N}{2} \times p$ matrix $\underset{\sim}{F}$ in (5.1.3) consists of the plus and minus ones in a two-level factorial design matrix. The element in the i^{th} $(1 \le i \le p)$ column of the matrix $\underset{\sim}{F}$ is the i^{th} coordinate of the design point in the p-dimensional design space. For the moment, we have assumed that $\frac{N}{2} = 2^p$. In Appendix I, the use of fractional factorials in the two halves of $\underset{\sim}{D}_w$ for reducing N, the number of experiments in (5.1.3), is discussed. We shall also from time to time refer to c as the "radius multiplier" since the design points in $\underset{\sim}{D}_w$, defined by (5.1.3), lie on a sphere of radius $c\sqrt{p}$.

The next two sections will be directed at the discussion of the first two considerations for choosing a first-degree design listed at the beginning of this section. Designs which are appropriate when neither the variance nor the bias error can be ignored are discussed in Appendix G,

as we said before. The third consideration (iii) is satisfied by the types of first-degree designs suggested. This point will be verified when we discuss second-degree designs in Chapter 7.

5.2 Designs to minimize average squared bias when the variance is assumed to be negligible.

As is shown in Section 4.2, when we fit a first-degree model and the true polynomial function is of degree 2, a source of error enters into the estimates of the regression coefficients. That is, bias enters into the first-degree estimates when it is impossible to represent the true function by an equation of the type fitted, a first-degree polynomial.

If we proceed as in Section 4.2, we want to construct a design such that the moment matrix of the design points equals the matrix of region moments. When the region of interest is R*, then from equation (4.2.10), the region moment matrix is

$$\underset{\sim}{\mu}_{11} = \Omega \int_{R^*} \underset{\sim}{w}_1 \underset{\sim}{w}_1' \, d\underset{\sim}{w} = \begin{bmatrix} 1 & \underset{\sim}{0}' & \pm 1 \\ \underset{\sim}{0} & \frac{\rho^{*2}}{p+2} \underset{\sim}{I}_p & \underset{\sim}{0} \\ \pm 1 & \underset{\sim}{0}' & 1 \end{bmatrix}. \quad (5.2.1)$$

Now for a first-degree rotatable design where the matrix $\underset{\sim}{W}_1$ is defined by (5.1.2), the moment matrix of the design points (see Appendix D) is of the form

$$\frac{1}{N}\underset{\sim}{W}_1'\underset{\sim}{W}_1 = \begin{bmatrix} 1 & \underset{\sim}{0}' & 0 \\ \underset{\sim}{0} & \lambda_2\underset{\sim}{I}_p & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & 1 \end{bmatrix}. \qquad (5.2.2)$$

In (5.2.1) and (5.2.2), $\underset{\sim}{I}_p$ is an identity matrix of order p. In equation (4.2.20) for B, we stated that if we set $\underset{\sim}{A} = (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{W}_2 = \underset{\sim}{\mu}_{11}^{-1}\underset{\sim}{\mu}_{12}$, i.e., $\frac{1}{N}\underset{\sim}{W}_1'\underset{\sim}{W}_1 = \underset{\sim}{\mu}_{11}$ and $\frac{1}{N}\underset{\sim}{W}_1'\underset{\sim}{W}_2 = \underset{\sim}{\mu}_{12}$, then B is minimized. However, when $\underset{\sim}{\mu}_{11}$ is defined as in (5.2.1), $\underset{\sim}{\mu}_{11}^{-1}$ does not exist since $\underset{\sim}{\mu}_{11}$ is of rank p+1. In Appendix G, it is shown that the average squared bias B is unaffected by the categorization of the components and depends only on λ_2, ρ^{*2}, p and $\tilde{\alpha}_{ij}$ (the standardized quadratic effects). Therefore, to say that $\underset{\sim}{\mu}_{11} = \frac{1}{N}\underset{\sim}{W}_1'\underset{\sim}{W}_1$, we are saying only that the upper left-hand (p+1) x (p+1) matrices in (5.2.1) and (5.2.2) are equal.

In setting the upper left-hand (p+1) x (p+1) matrices in (5.2.1) and (5.2.2) equal to one another, the following must be satisfied,

$$\lambda_2 = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^2 = \frac{\rho^{*2}}{p+2} \quad (1 \le i \le p), \qquad (5.2.3)$$

where ρ^* is the radius of the extended region of interest. If we use a design matrix of the form (5.1.3), i.e.,

$$\underset{\sim}{D}_w = c\begin{bmatrix} \underset{\sim}{F} \\ -\underset{\sim}{F} \end{bmatrix}$$

so that $\lambda_2 = c^2$, then by putting

$$c = \frac{\rho^*}{\sqrt{p+2}}, \qquad (5.2.4)$$

the equality in (5.2.3) is satisfied. Therefore, we see that to minimize bias alone, we want an orthogonal design with radius

$$\left(\frac{p}{p+2}\right)^{\frac{1}{2}} \rho^*. \tag{5.2.5}$$

In this chapter, we have been talking about first-degree designs in which the bottom half of the design matrix is the negative replicate of the top half of the design matrix. This property of the design matrix insured us that the third-degree moments of the designs would be zero. In addition, by using this particular form of the design matrix, we see that the estimates of the parameters (with the exception of the "mean" parameter) in the first-degree model will also be completely free of bias from the second-degree terms. We now attempt to clarify this concept.

Suppose with p independent variables we construct either the double-simplex or scaled two-level factorial design and denote the top half of $\underset{\sim}{D}_w$ by $\underset{\sim}{D}_{w1}$. In the simplex design, N_1 in $\underset{\sim}{D}_{w1}$ equals p+1 and with the two-level factorial, $N_1 = 2^p$ or some fraction thereof. If we let

$$\underset{\sim}{D}_w = \begin{bmatrix} \underset{\sim}{D}_{w1} \\ -\underset{\sim}{D}_{w1} \end{bmatrix}, \tag{5.2.6}$$

then the properties of rotatability are preserved, that is, the moments of degree 1 are zero and the moments of degree 2 are equal. Let the matrix of variables for the second-degree terms obtained by multiplying the elements of the columns of

$\underset{\sim}{D}_w$ in (5.2.6) be of the form

$$\begin{bmatrix} \underset{\sim}{D} \\ \underset{\sim}{D} \end{bmatrix},$$

for example, when p=2 and

$$\underset{\sim}{D}_{w1} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ -1 & 1 \\ 1 & 1 \end{bmatrix},$$

then

$$\underset{\sim}{D} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & 1 \end{bmatrix}.$$

The matrix of third-degree moments is null since

$$[\underset{\sim}{D}'_{w1}, -\underset{\sim}{D}'_{w1}] \begin{bmatrix} \underset{\sim}{D} \\ \underset{\sim}{D} \end{bmatrix} = [\underset{\sim}{D}'_{w1}\underset{\sim}{D} - \underset{\sim}{D}'_{w1}\underset{\sim}{D}] = \underset{\sim}{0}. \qquad (5.2.7)$$

The fact that all the third-degree moments are zero is important since these moments define all the elements except one row of the alias matrix. This matrix was denoted by $\underset{\sim}{A}$ in Section 4.2 and defined by

$$\underset{\sim}{A} = (\underset{\sim}{W}'_1\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}'_1\underset{\sim}{W}_2 . \qquad (5.2.8)$$

In (5.2.8), $\underset{\sim}{W}_1$ is the matrix of independent variables used in fitting the first-degree model and the matrix $\underset{\sim}{W}_2$ contains terms which are contained in the true polynomial function but are ignored in the assumed (fitted) model. We shall try to clarify this concept of bias expressed in (5.2.7) and (5.2.8) now with an example.

Double-simplex.

Let p=2 and define the first-degree model by

$$y = \alpha_0 + \alpha_1 w_1 + \alpha_2 w_2 + \alpha_{L_1} z_1 + \varepsilon.$$

In order for the second-degree model to contain all terms up to and including those of degree 2, we define the second-degree (true polynomial) model by

$$y = \alpha_0 + \alpha_1 w_1 + \alpha_2 w_2 + \alpha_{L_1} z_1 + \alpha_{11} w_1^2 + \alpha_{22} w_2^2 + \alpha_{12} w_1 w_2 + \varepsilon.$$

In matrix notation, the first-degree model is

$$\underset{\sim}{y} = \underset{\sim}{W}_1 \underset{\sim}{\alpha}_1 + \underset{\sim}{\varepsilon} \tag{5.2.9}$$

where $\underset{\sim}{\alpha}_1' = (\alpha_0, \alpha_1, \alpha_2, \alpha_{L_1})$ and a row of $\underset{\sim}{W}_1$ is $\underset{\sim}{w}_{u1}' = (1, w_{u1}, w_{u2}, z_{1u})$. The second-degree model in matrix notation is

$$\underset{\sim}{y} = \underset{\sim}{W}_1 \underset{\sim}{\alpha}_1 + \underset{\sim}{W}_2 \underset{\sim}{\alpha}_2 + \underset{\sim}{\varepsilon} \tag{5.2.10}$$

where $\underset{\sim}{\alpha}_2' = (\alpha_{11}, \alpha_{22}, \alpha_{12})$ and a row of $\underset{\sim}{W}_2$ is $\underset{\sim}{w}_{u2}' = (w_{u1}^2, w_{u2}^2, w_{u1}w_{u2})$. Now the matrices $\underset{\sim}{W}_1$ and $\underset{\sim}{W}_2$ corresponding to the above two models can be defined by (we shall not be concerned with minimizing the average squared bias but use a design matrix that is of general form),

$$\underset{\sim}{W}_1 = \begin{bmatrix} 1 & \rho & 0 & 1 \\ 1 & -\rho/2 & \sqrt{3}\rho/2 & 1 \\ 1 & -\rho/2 & -\sqrt{3}\rho/2 & 1 \\ 1 & -\rho & 0 & -1 \\ 1 & \rho/2 & -\sqrt{3}\rho/2 & -1 \\ 1 & \rho/2 & \sqrt{3}\rho/2 & -1 \end{bmatrix}, \quad \underset{\sim}{W}_2 = \begin{bmatrix} \rho^2 & 0 & 0 \\ \rho^2/4 & 3\rho^2/4 & -\sqrt{3}\rho^2/4 \\ \rho^2/4 & 3\rho^2/4 & \sqrt{3}\rho^2/4 \\ \rho^2 & 0 & 0 \\ \rho^2/4 & 3\rho^2/4 & -\sqrt{3}\rho^2/4 \\ \rho^2/4 & 3\rho^2/4 & \sqrt{3}\rho^2/4 \end{bmatrix}, \tag{5.2.11}$$

where ρ is the radius of the design. Then,

$$(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1} = \begin{bmatrix} 1/6 & & & 0 \\ & 1/(3\rho^2) & & \\ & & 1/(3\rho^2) & \\ 0 & & & 1/6 \end{bmatrix}, \quad \underset{\sim}{W}_1'\underset{\sim}{W}_2 = \begin{bmatrix} 3\rho^2 & 3\rho^2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

and the alias matrix is

$$\underset{\sim}{A} = (\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{W}_2 = \begin{bmatrix} \frac{\rho^2}{2} & \frac{\rho^2}{2} & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} . \qquad (5.2.12)$$

If $\underset{\sim}{\alpha}_1$ and $\underset{\sim}{\alpha}_2$ are defined as in (5.2.9) and (5.2.10), then from Section 4.2, the expected value of the estimates of the parameters in the first-degree model are

$$E(\hat{\underset{\sim}{\alpha}}_1) = \underset{\sim}{\alpha}_1 + \underset{\sim}{A}\,\underset{\sim}{\alpha}_2. \qquad (5.2.13)$$

Substituting $\underset{\sim}{A}$ and the vectors of parameters in (5.2.13), we have

$$\begin{aligned} E(\hat{\alpha}_0) &= \alpha_0 + \frac{\rho^2}{2}(\alpha_{11} + \alpha_{22}) \\ E(\hat{\alpha}_1) &= \alpha_1 \\ E(\hat{\alpha}_2) &= \alpha_2 \\ E(\hat{\alpha}_{L_1}) &= \alpha_{L_1} \quad . \end{aligned} \qquad (5.2.14)$$

Thus only the "mean" parameter is biased by second-degree effects, and the smaller the design radius ρ, the smaller the amount of bias.

Upon observing the matrix $\underset{\sim}{W}_1$ in (5.2.11), we see that we can rotate the design points say through a 45° angle such that the matrix $\underset{\sim}{W}_1$ is now

$$\underset{\sim}{W}_1 = \begin{bmatrix} 1 & \sqrt{3}\rho/2 & \rho/2 & 1 \\ 1 & 0 & -\rho & 1 \\ 1 & -\sqrt{3}\rho/2 & \rho/2 & 1 \\ 1 & -\sqrt{3}\rho/2 & -\rho/2 & -1 \\ 1 & 0 & \rho & -1 \\ 1 & \sqrt{3}\rho/2 & \rho/2 & -1 \end{bmatrix} \qquad (5.2.15)$$

and the $\underset{\sim}{W}_1'\underset{\sim}{W}_1$ matrix is the same as before. Because of this property of simplices, i.e., under an orthogonal transformation the pattern of design points goes over into a congruent pattern, it is difficult to talk specifically about the elements of the design matrix of simplex designs. Hence our reason for referring to the scaled two-level factorial in most of the material in this chapter. Also, it is interesting to note that if all the elements in the design matrix in (5.2.15) are multiplied by $c=\sqrt{p}/\sqrt{p+2}$, the minimum average square bias criterion is satisfied.

5.3 Designs to minimize average variance when the bias of the model is assumed to be zero.

As was shown in Section 4.2, equation (4.2.18), in order to minimize the quantity

$$V = \frac{N\Omega}{\sigma^2} \int_{R^*} \text{Var } \hat{y}(\underset{\sim}{w})\, d\underset{\sim}{w},$$

we want a design which minimizes the quantity

$$V = N[\text{Var } \hat{\alpha}_0 + \frac{\rho^{*2}}{p+2} \sum_{i=1}^{p} \text{Var } \hat{\alpha}_i + \text{Var } \hat{\alpha}_{L_1}].$$

Now with respect to the type of design, we know that V is minimized when $\underset{\sim}{W}_1$ is orthogonal. In terms of the size of the design however, if we look at the quantity

$$V = \text{trace } \{N(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{\mu}_{11}\} \qquad (5.3.1)$$

where

$$N(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1} = \begin{bmatrix} 1 & \underset{\sim}{0}' & 0 \\ \underset{\sim}{0} & \frac{1}{\lambda_2}\underset{\sim}{I}_p & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & 1 \end{bmatrix}, \text{ and } \underset{\sim}{\mu}_{11} = \begin{bmatrix} 1 & \underset{\sim}{0}' & \pm 1 \\ \underset{\sim}{0} & \frac{\rho^{*2}}{p+2}\underset{\sim}{I}_p & \underset{\sim}{0} \\ \pm 1 & 0' & 1 \end{bmatrix},$$

then V in (5.3.1) can be expressed as

$$V = \left(\frac{p}{p+2}\right)\frac{\rho^{*2}}{\lambda_2} + 2. \qquad (5.3.2)$$

In (5.3.2), since ρ^{*2} is fixed, we see that to minimize V, the largest λ_2 must be selected. This is achieved if all the design points lie on the boundary of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit in the factor space. Therefore, if the matrix $\underset{\sim}{D}_w$ is specified by (5.1.3), i.e.,

$$\underset{\sim}{D}_w = c\begin{bmatrix} \underset{\sim}{F} \\ -\underset{\sim}{F} \end{bmatrix},$$

and if we set

$$c = \frac{\rho^*}{\sqrt{p}}, \qquad (5.3.3)$$

then the average variance criterion (5.3.2) is minimized.

5.4 An example showing the design matrices of the mixture components corresponding to the "all-variance" and "all-bias" designs respectively.

Suppose the components x_1, x_2 belong to category 1 and x_3, x_4 belong to category 2. Also let $\underset{\sim}{x}_0'=(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$ and $\underset{\sim}{H}=\text{diag}(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Then from Section 4.3,

$$\rho^* = \frac{1}{4}\left\{\frac{1}{\frac{1}{16}} + \frac{1}{\frac{1}{16}}\right\}^{\frac{1}{2}} = \sqrt{2}.$$

Let us use the largest sphere R* as the region of interest and also use a design matrix of the form

$$\underset{\sim}{D}_w = c \begin{bmatrix} -1 & 1 \\ 1 & 1 \\ 1 & -1 \\ -1 & -1 \end{bmatrix} .$$

This design matrix corresponds to $\underset{\sim}{D}_{w1}$ in (5.2.6). Then if the matrix $\underset{\sim}{T}_1$ is of the form

$$\underset{\sim}{T}_1 = \begin{bmatrix} .707 & 0 \\ -.707 & 0 \\ 0 & .707 \\ 0 & -.707 \end{bmatrix} ,$$

and if we use $\underline{c}$ in (5.3.3), the levels of the mixture components corresponding to the "all-variance" design are

$$\underset{\sim}{D}_x = \begin{array}{c} \begin{matrix} \underset{\sim}{x}_1 & \underset{\sim}{x}_2 & \underset{\sim}{x}_3 & \underset{\sim}{x}_4 \end{matrix} \\ \begin{bmatrix} .073 & .427 & .427 & .073 \\ .427 & .073 & .427 & .073 \\ .427 & .073 & .073 & .427 \\ .073 & .427 & .073 & .427 \end{bmatrix} \end{array} . \qquad (5.4.1)$$

Now, if we use $\underline{c}$ in (5.2.4), the levels of the mixture components corresponding to the "all-bias" design are

$$\underset{\sim}{D}_x = \begin{bmatrix} .125 & .375 & .375 & .125 \\ .375 & .125 & .375 & .125 \\ .375 & .125 & .125 & .375 \\ .125 & .375 & .125 & .375 \end{bmatrix} . \qquad (5.4.2)$$

The design matrices (5.4.1) and (5.4.2) are found from the equation (3.2.15), i.e.,

$$\underset{\sim}{D}_x = \underset{\sim}{H}\, \underset{\sim}{\tilde{W}}\, \underset{\sim}{T}_1' + \underset{\sim}{x}_0\, \underset{\sim}{j}_4' \qquad (5.4.3)$$

where $\underset{\sim}{\tilde{W}} = \underset{\sim}{D}_w$.

VI. AN EXTENSION TO THREE AND MORE CATEGORIES.

In this chapter, we shall be concerned with extending the development of the mixture problem to cover situations where the number of categories of mixture components is greater than 2. In some sections of this chapter, the number of categories is general. In other sections, the theory is directed to situations involving just three and four categories of components. This limitation on the number of categories should not present any objections since in actual practice, it seems unlikely one would find applications of the mixture problem where $q \geq 5$. If however the procedure for $q>4$ is desired, the development is an obvious extension of that where $q \leq 4$. Also, in this chapter the discussion in the sections will be held to a minimum since most of the material is merely an extension of the theory developed for the two-category mixture problem.

6.1 General discussion concerning three categories.

In Section 1.3, we defined the mixtures stating that each mixture must be the result of some combination of at least one of the components from each of the categories. It was also mentioned that the weights assigned to each of the categories in defining the mixtures are equal. That is, if $q = 3$ and we let

$$S_\ell = \sum_{i=1}^{\ell} n_i \qquad (1 \leq \ell \leq 3) \qquad (6.1.1)$$

where in (6.1.1), n_i $(1 \le i \le 3)$ is the number of components in the i^{th} category, then

$$\sum_{j=S_{\ell-1}+1}^{S_\ell} x_j = \frac{1}{3} \qquad (1 \le \ell \le 3) \tag{6.1.2}$$

and

$$\sum_{\ell=1}^{3} \sum_{j=S_{\ell-1}+1}^{S_\ell} x_j = 1. \tag{6.1.3}$$

The shape (number of extremities) of the convex polytope again depends on the number of components in each of the categories. However, it is difficult to generalize the formula for the number of extremities (denoted by $\tilde{N}$ in Section 2.3) as we did for two categories. This is because of the increasing number of components with the additional number of categories (i.e., $k \ge 2q$). On the other hand, we tabulate $\tilde{N}$ (Table 1. on the next page) associated with the polytope for various n_i $(1 \le i \le 3)$ for the three-category case to indicate the complexity of the convex polytope or factor space of the mixture components.

To calculate the number $\tilde{N}$ of extremities of dimensionality $k-3-\mu$, we proceed as with two categories by defining μ $(1 \le \mu \le k-3)$ as the number of components set equal to zero. For example, let $\mu=2$, $n_1=2$, $n_2=2$ and $n_3=3$. Then

$$\tilde{N} = \binom{7}{2} - \left[\binom{2}{2} + \binom{2}{2}\right]$$

$$= 21 - 2 = 19.$$

Table I.

The Number of Extremities of the Polytope When the Components Come From Three Categories.

μ	n_1	n_2	n_3	$\tilde{N}$
	2	2	2	
1				6
2				12
3				8
	3	2	2	
1				7
2				19
3				24
4				12
	4	2	2	
1				8
2				26
3				44
4				40
5				16
	3	3	2	
1				8
2				27
3				48
4				45
5				18

μ	n_1	n_2	n_3	$\tilde{N}$
	3	3	3	
1				9
2				36
3				81
4				108
5				81
6				27
	4	3	2	
1				9
2				35
3				76
4				98
5				72
6				24

If $\mu=4$, $n_1=2$, $n_2=2$ and $n_3=3$, then

$$\tilde{N} = \binom{7}{4} - [\binom{n_1}{2}\binom{n_2}{1}\binom{n_3}{1} + \binom{n_1}{2}\binom{n_2}{2} + \binom{n_1}{2}\binom{n_3}{2}$$

$$+ \binom{n_2}{2}\binom{n_3}{2} + \binom{n_1}{1}\binom{n_2}{2}\binom{n_3}{1} + \binom{n_2}{1}\binom{n_3}{3} + \binom{n_1}{1}\binom{n_3}{3}]$$

$$= 35 - (6 + 1 + 3 + 3 + 6 + 2 + 2) = 12.$$

The values of $\tilde{N}$ in Table 1. can be checked using the generalization of Euler's formula,

$$\tilde{N}_0 - \tilde{N}_1 + \tilde{N}_2 - \ldots (-1)^{r-1}\tilde{N}_{r-1} = 1 - (-1)^r$$

where $\tilde{N}_i$ $(0\leq i\leq r-1)$ is the number of boundaries of dimensionality i of the r-dimensional polytope and $r=n_1+n_2+n_3-3$. For example, for $r=2+2+2-3=3$, then

$$8 - 12 + 6 = 2$$

$$= 1 - (-1)^3.$$

For $r=3+2+2-3=4$, then

$$12 - 24 + 19 - 7 = 0$$

$$= 1 - (-1)^4.$$

Now the general polynomial model (1.1.2) can again be defined so that the terms in the model represent valid mixture combinations of the components. The desirability of a representation of this form as stated in Section 2.2 is that the terms in the model can be put into a one to one correspondence with the design points suggested by this expression. Thus the estimated response at some points in the factor space can simply be represented by a single term in the model. For example, if we want a quadratic model representation and

$n_1 = n_2 = n_3 = 2$ when $q = 3$, then if we apply the constraints

$$x_1 + x_2 = x_3 + x_4 = x_5 + x_6 = \frac{1}{3} \qquad (6.1.4)$$

while at the same time repeatedly multiplying the terms in the model

$$y = \sum_{i=1}^{6} \beta_i x_i + \sum\sum_{i<j}^{6} \beta_{ij} x_i x_j + \epsilon$$

by $x_1+x_2+x_3+x_4+x_5+x_6=1$, the quadratic model can be written as

$$y = \sum_{i=1}^{2}\sum_{j=3}^{4}\sum_{\ell=5}^{6} \gamma_{ij\ell} x_i x_j x_\ell + \sum_{i=1}^{2}\sum_{j=3}^{4} \gamma_{ij56} x_i x_j x_5 x_6$$

$$+ \sum_{i=1}^{2}\sum_{\ell=5}^{6} \gamma_{i34\ell} x_i x_3 x_4 x_\ell + \sum_{j=3}^{4}\sum_{\ell=5}^{6} \gamma_{12j\ell} x_1 x_2 x_j x_\ell + \epsilon. \qquad (6.1.5)$$

The estimated response at the mixture combination $x_1=x_3=x_5=\frac{1}{3}$ is simply $\hat{y} = \frac{\hat{\gamma}_{135}}{9}$ where $\hat{\gamma}_{135}$ is the estimate of γ_{135}.

One disadvantage of the form (6.1.5) of the polynomial is that it contains an excessive number of terms (namely 20). Furthermore, without adding a center point or without replicating experiments, there is no way to measure the lack of fit of the model. This is because the model contains the same number of terms as there are design points suggested by this expression. The design space is a 3-dimensional cube and the design points lie at the vertices and midpoints of the edges. Therefore, we shall proceed as with two categories (Chapter 3) to transform the mixture components to independent variables so as to be able to define the problem with fewer variables.

6.2 Transformation to independent variables: Three and four categories.

To briefly review the steps in transforming to independent variables, it was explained in Chapter 3 that with q categories of k mixture components, the rank of the matrix $\underset{\sim}{V}$ in the general linear model $\underset{\sim}{y} = \underset{\sim}{V}\underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ is k-q. This led to the reparametrization of the model to a model $\underset{\sim}{y} = \underset{\sim}{\tilde{W}}\underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}$ where a row of the matrix $\underset{\sim}{\tilde{W}}$ contains the levels of the k-q independent variables. Now if we let q=3 and p=k-3, then the transformation from the variables v_j $(1 \le j \le k)$ to the design variables w_i $(1 \le i \le k-3)$, which was shown for the case q=2 in Section 3.2, extends as follows to the case where q=3,

$$\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_p \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} & & & & & & & & & & \\ & & & & & \underset{\sim}{T}_1' & & & & & \\ & & & & & & & & & & \\ \tilde{h}_1 & \tilde{h}_2 & \cdots & \tilde{h}_{S_1} & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & \tilde{h}_{S_1+1} & \cdots & \tilde{h}_{S_2} & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & \tilde{h}_{S_2+1} & \cdots & \tilde{h}_k \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ \\ \\ v_k \end{bmatrix} \qquad (6.2.1)$$

where the matrix $\underset{\sim}{T}_1'$ is (k-3) x k and

$$\tilde{h}_i = \frac{h_i}{h^{1*}} \qquad (1 \le i \le S_1)$$

$$\tilde{h}_j = \frac{h_j}{h^{2*}} \qquad (S_1+1 \le j \le S_2) \qquad (6.2.2)$$

$$\tilde{h}_\ell = \frac{h_\ell}{h^{3*}} \qquad (S_2+1 \le \ell \le k)$$

and from (6.1.1),

$$S_i = \sum_{\ell=1}^{i} n_\ell \quad , \quad h^{i'*} = \Big(\sum_{j=S_{i'-1}+1}^{S_{i'}} h_j^2 \Big)^{\frac{1}{2}} \quad i'=1,2,3. \tag{6.2.3}$$

The last three rows of the transformation matrix in (6.2.1) can be put in vector form by

$$\begin{aligned} \tilde{\underset{\sim}{h}}' &= (\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_{S_1}, 0, 0, \ldots, 0) \\ \bar{\underset{\sim}{h}}' &= (0, 0, \ldots, 0, \tilde{h}_{S_1+1}, \ldots, \tilde{h}_{S_2}, 0, \ldots, 0) \\ \hat{\underset{\sim}{h}}' &= (0, 0, \ldots, 0, 0, 0, \ldots, 0, \tilde{h}_{S_2+1}, \ldots, \tilde{h}_k), \end{aligned} \tag{6.2.4}$$

and therefore, (6.2.1) can be rewritten as

$$\begin{bmatrix} \underset{\sim}{w} \\ \underset{\sim}{0} \end{bmatrix} = \underset{\sim}{T}' \underset{\sim}{v} \tag{6.2.5}$$

where $\underset{\sim}{T} = [\underset{\sim}{T}_1 \; \tilde{\underset{\sim}{h}} \; \bar{\underset{\sim}{h}} \; \hat{\underset{\sim}{h}}]$. In (6.2.1) and (6.2.5), $\underset{\sim}{T}$ is defined to be a k x k orthogonal matrix and with the elements of the vectors in (6.2.4) specified by (6.2.2), the matrix $\underset{\sim}{T}_1$ can be constructed from the h_i ($1 \le i \le k$) so that the matrix $\underset{\sim}{T}$ is orthogonal (see Appendices E, F).

For the construction of the N x k matrix $\underset{\sim}{W}$ of independent variables, we stated in Section 3.3 that the design matrix will be divided into q compartments where each compartment contains N_j ($1 \le j \le q$) rows. When q=3 and the N_j are equal, the coefficients z_{iu} ($1 \le i \le 2$) of the contrast columns are determined by equations (3.3.6), (3.3.6A) and (3.3.8). The form of the matrix $\underset{\sim}{W}$ is

$$\underset{\sim}{W} = \begin{bmatrix} & & \underset{\sim}{1}_{N/3} & \underset{\sim}{1}_{N/3} \\ \underset{\sim}{j}_N & \underset{\sim}{D}_w & -\underset{\sim}{1}_{N/3} & \underset{\sim}{1}_{N/3} \\ & & \underset{\sim}{0}_{N/3} & -\underset{\sim}{2}_{N/3} \end{bmatrix} . \qquad (6.2.6)$$

Note that to form the matrix

$$[\underset{\sim}{0}\ \tilde{\underset{\sim}{W}}\ \underset{\sim}{0}\ \underset{\sim}{0}]$$

where $\underset{\sim}{0}$ is an N x 1 column vector of zeroes so that columns of elements (not all zero) can be added to form the matrix $\underset{\sim}{W}$ in (6.2.6), the columns of the matrix $\underset{\sim}{T}$ must be rearranged to

$$\underset{\sim}{T} = [\tilde{\underset{\sim}{h}}\ \underset{\sim}{T}_1\ \bar{\underset{\sim}{h}}\ \hat{\underset{\sim}{h}}] . \qquad (6.2.7)$$

Then over N observations, the transformation in (6.2.5) is

$$\underset{\sim}{V}\ \underset{\sim}{T} = [\underset{\sim}{0}\ \tilde{\underset{\sim}{W}}\ \underset{\sim}{0}\ \underset{\sim}{0}] . \qquad (6.2.7A)$$

When q=4 and the N_j $(1 \le j \le 4)$ are equal, an additional column is added to the matrix $\underset{\sim}{W}$ in (6.2.6) so that $\underset{\sim}{W}$ is now

$$\underset{\sim}{W} = \begin{bmatrix} & & \underset{\sim}{1}_{N/4} & \underset{\sim}{1}_{N/4} & \underset{\sim}{1}_{N/4} \\ \underset{\sim}{j}_N & \underset{\sim}{D}_w & -\underset{\sim}{1}_{N/4} & \underset{\sim}{1}_{N/4} & \underset{\sim}{1}_{N/4} \\ & & \underset{\sim}{0} & -\underset{\sim}{2}_{N/4} & \underset{\sim}{1}_{N/4} \\ & & \underset{\sim}{0} & \underset{\sim}{0} & -\underset{\sim}{3}_{N/4} \end{bmatrix} . \qquad (6.2.8)$$

To determine the settings of the levels of the mixture components to be performed experimentally, the procedure is identical to that of two categories. That is to say, if

$$\underset{\sim}{X} = [\tilde{\underset{\sim}{W}}\ \underset{\sim}{0}] \begin{bmatrix} \underset{\sim}{T}_1' \\ \underset{\sim}{T}_2' \end{bmatrix} \underset{\sim}{H} \qquad (6.2.9)$$

$$= \tilde{\underset{\sim}{W}}\ \underset{\sim}{T}_1'\ \underset{\sim}{H} ,$$

then

$$x_{uj} = \left(\sum_{\ell=1}^{k-q} w_{u\ell} t_{j\ell} \right) h_j + x_{0j} \qquad (6.2.10)$$

where $w_{u\ell}$ is the element in the u^{th} row and ℓ^{th} column ($1 \le u \le N$, $1 \le \ell \le k-q$) of the matrix $\tilde{\underset{\sim}{W}}$ and $t_{j\ell}$ is the element in the j^{th} row and ℓ^{th} column of $\underset{\sim}{T}$ ($1 \le j \le k$, $1 \le \ell \le k-q$) where the matrix $\underset{\sim}{T}$ is defined as in (6.2.5). Note that if the matrix $\tilde{\underset{\sim}{W}}$ is located in the matrix $\underset{\sim}{W}$ as in (6.2.7A), and the matrix $\underset{\sim}{T}$ is of the form (6.2.7), then

$$x_{uj} = \left(\sum_{\ell=2}^{k-q+1} w_{u\ell} t_{j\ell} \right) h_j + x_{0j} \qquad (1 \le j \le k,\ 2 \le \ell \le k-q+1).$$

6.3 Radius of the largest sphere: Three categories.

When q=3, the shortest distance from the point $\underset{\sim}{w}=\underset{\sim}{0}$ to an extremity of the polytope is equal to the length of the shortest orthogonal projection from $\underset{\sim}{w}=\underset{\sim}{0}$ to an extremity of dimensionality k-4. Thus the radius ρ^* of the largest sphere centered at the point $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the convex polytope is

$$\rho^* = \min_{1 \le i \le 3} \rho_i \qquad (6.3.1)$$

where

$$\rho_i = \min_{S_{i-1}+1 \le j \le S_i} x_{0j} \left\{ \frac{1}{h_j^2} + \frac{1}{\sum_{\substack{j'=S_{i-1}+1 \\ j' \ne j}}^{S_i} h_{j'}^2} \right\}^{\frac{1}{2}}, \qquad (6.3.2)$$

6.4 Design matrices for three and four categories.

When the mixture problem involved just two categories of components, the design matrix $\underset{\sim}{D}_w$ was constructed such that the levels of the design variables in the bottom half of $\underset{\sim}{D}_w$ are the negative of the levels of the variables in the top half of $\underset{\sim}{D}_w$. In addition, each half of the matrix $\underset{\sim}{D}_w$ consisted of a first-degree orthogonal design. This partitioning first of all enabled us to estimate the effects of all the terms in the first-degree model separately from one another. Secondly, we were able to keep the estimates of the first-degree effects free of bias from possible second-degree effects by assuring that the third-degree moments of the design vanish.

Now in order to maintain the above conditions of symmetry and orthogonality when q=3 and q=4, let us define the design matrices $\underset{\sim}{D}_w$ in the following manner. For three categories,

$$\underset{\sim}{D}_w = c\begin{bmatrix} \underset{\sim}{F} \\ \underset{\sim}{n} \\ -\underset{\sim}{F} \end{bmatrix} \qquad (6.4.1)$$

where with the simplex design, $\underset{\sim}{F}$ is a (p+1) x p matrix and the p+1 rows consist of design points which lie at the vertices of a p-dimensional simplex. For the scaled two-level factorial design, the ℓ x p matrix $\underset{\sim}{F}$ (where ℓ=N/3 if the sizes of the compartments in $\underset{\sim}{D}_w$ are equal) consists of the plus and minus ones in a two-level factorial design matrix and the elements of the p columns of $\underset{\sim}{F}$ represent the levels

of the variables in the first-degree model.

In (6.4.1), $\underset{\sim}{n}$ is a matrix consisting of center point replicates, i.e., $w_{ui}=0$ $(1\leq i\leq p)$. In (6.2.6), $\underset{\sim}{n}$ is N_2 x p where $N_2=N/3$ but this is not necessary. That is, if $N_1\neq N_2\neq N_3$ (where N_j is the number of rows in the j^{th} compartment of $\underset{\sim}{D}_w$), then the $z_{i.}^{(i+1)}$, the coefficients in the $(i+1)^{st}$ compartment corresponding to the i^{th} contrast $(1\leq i\leq 2)$, will be calculated by the formula (3.3.7). Finally, the total number of experiments with the simplex design is $N=2(p+1)+N_2$ and with the factorial design, $N=2\cdot 2^{p-i}+N_2$ where $i=0,1,2,\ldots$

For the case of four categories, the design matrix of the design variables will be

$$\underset{\sim}{D}_w = c\begin{bmatrix} \underset{\sim}{D}_{w1} \\ -\underset{\sim}{D}_{w1} \end{bmatrix} . \qquad (6.4.2)$$

With the simplex design,

$$\underset{\sim}{D}_{w1} = \begin{bmatrix} \underset{\sim}{F} \\ -\underset{\sim}{F} \end{bmatrix} \qquad (6.4.3)$$

or

$$\underset{\sim}{D}_{w1} = \begin{bmatrix} \underset{\sim}{F} \\ \underset{\sim}{F} \end{bmatrix} \qquad (6.4.4)$$

where $\underset{\sim}{F}$ is defined as in (6.4.1). The scaled two-level factorial will also have a design matrix of the form (6.4.2) where $\underset{\sim}{D}_{w1}=\underset{\sim}{F}$ but the elements of the N/2 x p matrix $\underset{\sim}{F}$ will be the plus and minus ones in a two-level factorial design matrix. With the simplex design, $N=4(p+1)$ and with the factorial design, $N=2\cdot 2^{p-i}$ where $i=0,1,2,\ldots$ Actually, the only requirements for the form (6.4.2) of the design matrix

is that the design matrix be divided into four parts where each part or compartment contains a first-degree orthogonal design, and the columns in each compartment must sum to zero.

6.5 Moment matrices for three and four categories: A comparison between orthogonal contrasts and orthogonal blocking.

We stated in Section 3.3 that we would try to justify our choice of using orthogonal contrasts rather than orthogonal blocking. One of the reasons for this choice is in the construction of the moment matrix. We now show the general forms of the moment matrices corresponding to both methods in an attempt to clarify this point.

Let the number N_j $(1 \le j \le q)$ of rows in the j compartments of $\underset{\sim}{D}_w$ be equal. This enables us to generalize the form of the matrix $\underset{\sim}{D}_w$. If q=3 and the matrices $\underset{\sim}{W}$ and $\underset{\sim}{D}_w$ are defined by (6.2.6) and (6.4.1) respectively, the moment matrix is

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' & 0 & 0 \\ \underset{\sim}{0} & \lambda_2 \underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & 2/3 & 0 \\ 0 & \underset{\sim}{0}' & 0 & 2 \end{bmatrix}, \qquad (6.5.1)$$

where

$$\lambda_2 = \sum_{u=1}^{N} \frac{w_{ui}^2}{N} \qquad (1 \le i \le p).$$

If q=4 and $\underset{\sim}{D}_W$ is defined by (6.4.2) and $\underset{\sim}{W}$ is defined as in Section 6.2, (6.2.8), then

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' & 0 & 0 & 0 \\ \underset{\sim}{0} & \lambda_2\underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & 1/2 & 0 & 0 \\ 0 & \underset{\sim}{0}' & 0 & 2/3 & 0 \\ 0 & \underset{\sim}{0}' & 0 & 0 & 3 \end{bmatrix}. \qquad (6.5.2)$$

By observing the matrices in (6.5.1) and (6.5.2) above, we see that the general form of the moment matrix with q categories using orthogonal contrasts (where the N_j $(1 \leq j \leq q)$ are equal) is

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' & \underset{\sim}{0} \\ \underset{\sim}{0} & \lambda_2\underset{\sim}{I}_p & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0}' & \underset{\sim}{D} \end{bmatrix} \qquad (6.5.3)$$

where the matrix $\underset{\sim}{D}$ is a (q-1) x (q-1) diagonal matrix and the diagonal elements are equal to $\frac{2}{q}$, $\frac{6}{q}$, $\frac{12}{q}$, ..., q. When the N_j $(1 \leq j \leq q)$ are not equal, $\underset{\sim}{D}$ is still diagonal but the elements are different from those in (6.5.3).

If the method of orthogonal blocking is used to measure the variation between the groups of observations, the formulas used for calculating the coefficients in the block contrasts are different from the formulas (3.3.7) or (3.3.8). In some cases, the calculations of the coefficients in the block contrasts are considerably more time consuming. Also, although the form of the moment matrix using orthogonal blocking is the same as (6.5.3), the matrix $\underset{\sim}{D}$ in (6.5.3) will

not be diagonal. Instead, using orthogonal blocking where the N_j are equal,

$$\underset{\sim}{D} = \left[\frac{1}{q}\, \underset{\sim}{I}_{q-1} - \underset{\sim}{j}\, \underset{\sim}{j}'\right] \qquad (6.5.4)$$

where $\underset{\sim}{j}$ is a (q-1) x 1 column vector of ones.

Considering the forms of the moment matrices using both methods, we see that one distinct advantage in using orthogonal contrasts is in obtaining the precision matrix $N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}$ (the inverse of the moment matrix). This inverse matrix is easily obtained because of the diagonal property of the moment matrix. Besides obtaining estimates of the parameters in the model, the precision matrix is also used for comparing the efficiencies of designs (although the block variables are of no concern). Therefore, one might choose to use the method of orthogonal contrasts to that of orthogonal blocking simply because of the ease in inverting a diagonal matrix.

Actually this ease in calculations may be debatable since the matrix $\underset{\sim}{D}$ in (6.5.4) is of the form

$$\underset{\sim}{D} = \begin{bmatrix} a & b & b & \cdots & b \\ b & a & b & \cdots & b \\ \vdots & & \ddots & & b \\ b & b & \cdots & b & a \end{bmatrix}.$$

The inverse $\underset{\sim}{D}^{-1}$ is of the same form and requires only the solution of two equations in the two unknowns a and b. That is to say, the inverse $\underset{\sim}{D}^{-1}$ will be of the form

$$\underset{\sim}{D}^{-1} = \begin{bmatrix} c & d & d & \cdots & d \\ d & c & d & \cdots & d \\ d & & \ddots & & d \\ \vdots & & & \ddots & \vdots \\ d & d & \cdots & d & c \end{bmatrix}$$

where

$$c = \frac{a+(q-3)b}{a^2+(q-3)ab-(q-2)b^2} \quad , \quad d = \frac{-b}{a^2+(q-3)ab-(q-2)b^2} \, .$$

When $a = \frac{1}{q} - 1$ and $b = -1$, then

$$c = \frac{q(1+2q-q^2)}{1+q-q^2} \quad , \quad d = \frac{q^2}{1+q-q^2} \, .$$

In any event, the reader now has an alternative method from orthogonal blocking for extracting information when blocking the observations. In the next chapter, we extend the method of orthogonal contrasts in covering second-degree designs.

6.6 The effect of the addition of categories on the properties of $\hat{y}$.

The average squared bias was defined in Section 4.2 by

$$B = \frac{N}{\sigma^2} \underset{\sim}{\alpha}_2' \underset{\sim}{A} \underset{\sim}{\alpha}_2 \, . \qquad (6.6.1)$$

Now since the contrast terms in the model are the only terms affected by the addition of categories of mixture components and these terms do not have any influence on the fit of the model to the physical situation, it can easily be shown that B in (6.6.1) is unaffected by the addition of categories of components. Since the quantity B involves only the quantities λ_2, ρ^{*2}, p and $\tilde{\alpha}_{ij}$ (the standardized quadratic effects

of the true surface), we shall omit the details of proving the above statement.

The average variance criterion, on the other hand, is affected by the inclusion of additional terms ($k \geq 2q$) to the model. That is to say, from Chapter 4, equations (4.2.17) and (4.2.10),

$$V = \text{trace}\ \{N(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{\mu}_{11}\}$$

$$= \left(\frac{p}{p+2}\right)\frac{\rho^{*2}}{\lambda_2} + K, \qquad (6.6.2)$$

where the quantity K depends on the contrast terms in the model. The orders of the matrices $N(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}$ and $\underset{\sim}{\mu}_{11}$ are increased by the addition of categories which inturn affects K in (6.6.2). However, the procedure used to minimize V does not change in that we want to use a design with a large λ_2 since ρ^{*2} is fixed. This means that we want to use a design whose points all lie on the boundary of the largest p-dimensional sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the factor space (convex polytope).

VII. SECOND-DEGREE DESIGNS.

In Chapter 5, designs used when fitting a first-degree linear model were introduced. Within the region of experimentation however, there often exists some degree of curvature of the response surface. When curvature exists and it is desired to obtain a measure of this curvature, then generally the fitting of only a first-degree model is not sufficient. That is to say, in order to obtain some measure of the curvature of the response surface, we must fit a model which is at least of the second degree.

In this chapter therefore, we shall direct our attention to the construction of second-degree designs and the associated second-degree model in the design variables w_i $(1 \leq i \leq k-q)$. The second-degree designs suggested for use will possess the property that they are easily formed by augmenting the previously constructed first-degree designs. In this chapter, we shall also extend the development of the method of orthogonal contrasts so as to be able to measure the variation that might exist between stages of experimentation where a stage corresponds to any condition which is thought to be homogeneous such as a period of time, a particular piece of production machinery, a chemist, etc. The analysis of a second-degree central composite design, that is, the calculation of the estimates of the parameters in the second-degree model using a central composite design

is shown in Appendix H.

7.1 The second-degree model in the design variables and the general form of the moment matrix for a second-degree design.

Let the number of categories of mixture components be two (i.e., q=2). Then the second-degree model in the design variables can be written as

$$y = \underset{\sim}{w}_1'\underset{\sim}{\alpha}_1 + \underset{\sim}{w}_2'\underset{\sim}{\alpha}_2 + \epsilon \qquad (7.1.1)$$

where $\underset{\sim}{w}_1' = (1, w_1, \ldots, w_{k-2}, \pm 1)$ and

$$\underset{\sim}{w}_2' = (w_1^2, \ldots, w_{k-2}^2, \sqrt{2}w_1w_2, \ldots, \sqrt{2}w_{k-3}w_{k-2}).$$

The coefficient $\sqrt{2}$ of the crossproduct terms in $\underset{\sim}{w}_2'$ is standard notation in response surface work and will not complicate the analysis as we shall see. In the model (7.1.1), the vectors of parameters are

$$\underset{\sim}{\alpha}_1' = (\alpha_0, \alpha_1, \ldots, \alpha_{k-2}, \alpha_{L_1})$$

and

$$\underset{\sim}{\alpha}_2' = (\alpha_{11}, \ldots, \alpha_{k-2,k-2}, \frac{\alpha_{12}}{\sqrt{2}}, \ldots, \frac{\alpha_{k-3,k-2}}{\sqrt{2}}). \qquad (7.1.2)$$

Over all N observations, the model will be written as

$$\underset{\sim}{y} = \underset{\sim}{W}\,\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon} \qquad (7.1.3)$$

where the u^{th} row of the matrix $\underset{\sim}{W}$ is $\underset{\sim}{w}_u' = (\underset{\sim}{w}_{1u}', \underset{\sim}{w}_{2u}')$ and $\underset{\sim}{\alpha}' = (\underset{\sim}{\alpha}_1', \underset{\sim}{\alpha}_2')$. Note that the matrix $\underset{\sim}{W}$ was used previously to denote the matrix of independent variables but in this chapter, when we refer to the matrix $\underset{\sim}{W}$, we shall always mean the form used in (7.1.3).

To determine the form of the moment matrix, for general q, of a second-degree design, let us rearrange the contrast column in the matrix $\underset{\sim}{W}$ as well as the contrast parameter in $\underset{\sim}{\alpha}$ and let p=k-q. Then the matrix $\underset{\sim}{W}$ for general q is

$$\underset{\sim}{W} = \begin{bmatrix} 1, w_1, \ldots, w_p, w_1^2, \ldots, w_p^2, \sqrt{2}w_1w_2, \ldots, \sqrt{2}w_{p-1}w_p, z_1, \ldots, z_{q-1} \\ \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots \end{bmatrix} \tag{7.1.4}$$

where the elements in the z_i $(1 \le i \le q-1)$ columns take the values of the z_{ui} $(1 \le i \le q-1)$ in the equations (3.3.6), (3.3.6A) and (3.3.7). The general form of the moment matrix for any second-degree design where the matrix $\underset{\sim}{W}$ is defined by (7.1.4) is

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' & \lambda_2\underset{\sim}{j}_p' & \underset{\sim}{0}' & \underset{\sim}{0}' & 0 \\ \underset{\sim}{0} & \lambda_2\underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \lambda_2\underset{\sim}{j}_p & \underset{\sim}{0} & \begin{matrix} G & \lambda_4 & \cdots & \lambda_4 \\ \lambda_4 & G & \cdots & \lambda_4 \\ \vdots & & \ddots & \vdots \\ \lambda_4 & \lambda_4 & \cdots & G \end{matrix} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{P} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & 2\lambda_4\underset{\sim}{I}_{\binom{p}{2}} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & \underset{\sim}{P}' & \underset{\sim}{0}' & \underset{\sim}{0}' & d \end{bmatrix}, \tag{7.1.5}$$

where the quantities in the moment matrix are

$$\lambda_2 = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^2 \qquad (1\leq i\leq p)$$

$$\lambda_4 = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^2 w_{uj}^2 \qquad (1\leq i<j\leq p)$$

$$G = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^4 \qquad (1\leq i\leq p) \qquad (7.1.6)$$

$$N\underset{\sim}{Q} = \text{diag}\left(\sum_{u=1}^{N} z_{iu}^2\right) \qquad (1\leq i\leq q-2)$$

$$P_i = \frac{1}{N}\sum_{u=1}^{N} z_{q-1,u} w_{ui}^2 \qquad (1\leq i\leq p)$$

$$d = \frac{1}{N}\sum_{u=1}^{N} z_{q-1,u}^2 \quad .$$

Note that the moments λ_2, λ_4 and G are independent of the i ($1\leq i\leq p$). The quantity P_i in (7.1.6) will be discussed later in Section 7.3.

We shall now discuss the construction of second-degree designs and also explain how the quantities in (7.1.6) influence the strategy of the experimenter in choosing a second-degree design. In Chapters 5 and 6, the discussion was limited to rotatable designs and we shall maintain this requirement of rotatability for the second-degree designs discussed in this chapter.

7.2 Second-degree rotatable arrangements and designs.

Two of the most commonly used second-degree rotatable designs are the equiradial rotatable designs and the rotatable composite designs. In a manner of speaking, the rotatable composite designs belong to the class of equiradial designs. When however we speak of equiradial designs, we are referring to a class of designs in which the design points are equi-spaced on the boundaries of two or more concentric p-dimensional spheres and the radii of the spheres are unequal and greater than zero. Also, the number of points on the boundary of each sphere, whose radius is greater than zero, must be greater than or equal to five. In contrast to this, with the rotatable central composite designs, one of the spheres has radius equal to zero.

A set of design points is said to form a rotatable arrangement of the second degree in p variables if

$$\sum_{u=1}^{N} w_{ui}^2 = N \lambda_2 \qquad (1 \le i \le p) \qquad (7.2.1)$$

$$\sum_{u=1}^{N} w_{ui}^4 = 3N\lambda_4 \qquad (1 \le i \le p)$$

$$= \sum_{u=1}^{N} 3w_{ui}^2 w_{uj}^2 \qquad (1 \le i < j \le p), \qquad (7.2.2)$$

and all other sums of powers and products up to and including degree 4 are zero (see Appendix D). Now the point set is said to form a rotatable design if the above equations

(7.2.1) and (7.2.2) are satisfied and the moment matrix (7.1.5) is non-singular. Box and Hunter [6] show that a necessary and sufficient condition for the moment matrix to be non-singular is that

$$\frac{\lambda_4}{\lambda_2^2} > \frac{p}{p+2}, \tag{7.2.3}$$

a condition which is always satisfied, as pointed out in [6], merely by the addition of center point replicates to the design. The inequality (7.2.3) becomes an equality only when all of the design points lie on a p-dimensional sphere and consequently the moment matrix is singular. This is seen by the following example.

Let ρ be the common distance from the design points to the center of the design for a single set of design points. Then, we can see that in the $\underset{\sim}{W}$ matrix (7.1.4),

$$\sum_{u=1}^{p} w_{ui}^2 = \rho^2 \, w_{u0} \qquad (1 \le u \le N) \tag{7.2.4}$$

where $w_{u0}=1$, and therefore, the matrix $\underset{\sim}{W}$ is deficient. Consequently, the constant parameter α_0 in the model (7.1.3) cannot be separately estimated from the quadratic parameters α_{ii} $(1 \le i \le p)$. Now, if we have just one single set of design points, then from (7.2.4),

$$\begin{aligned} \rho^2 &= \frac{1}{N} \sum_{u=1}^{N} \sum_{i=1}^{p} w_{ui}^2 \\ &= p\lambda_2. \end{aligned} \tag{7.2.5}$$

If we let

$$\rho^4 = (\rho^2)^2 = \frac{1}{N}\sum_{u=1}^{N}\Big(\sum_{i=1}^{p} w_{ui}^2\Big)^2 ,$$

then

$$\rho^4 = \sum_{i=1}^{p}\frac{1}{N}\sum_{u=1}^{N} w_{ui}^4 + \sum_{i=1}^{p}\sum_{j\neq i}^{p}\frac{1}{N}\sum_{u=1}^{N} w_{ui}^2 w_{uj}^2$$

$$= 3p\lambda_4 + p(p-1)\lambda_4. \qquad (7.2.6)$$

But since $\rho^4=(\rho^2)^2$, that is,

$$p^2\lambda_2^2 = \lambda_4(p^2 + 2p),$$

then

$$\frac{\lambda_4}{\lambda_2^2} = \frac{p}{p+2} . \qquad (7.2.7)$$

A result of the equality (7.2.7) as we shall see later in Section 7.4, is that the moment matrix associated with the set of points only on the boundary of one sphere is singular and therefore, the estimates of the parameters cannot be found by the usual least-squares procedure.

In the construction of a second-degree design, we have shown that if the experimenter places a set of n_1 points at a distance $\rho_1\neq 0$ from the center of the design, he must augment the design with an additional set of n_2 points at a distance $\rho_2\neq\rho_1$ from the center or at the center. Of course, the quantities λ_2 and λ_4 (see (7.1.6)) in the moment matrix will be affected by the pattern of design points and this

the experimenter must consider prior to constructing the design.

One further point to consider prior to constructing second-degree designs is the dimensionality of the factor space (number of independent variables). In two dimensions, there are an infinite number of concentric equiradial sets of points but in using regular figures, the pentagonal and hexagonal arrangements each with center point replicates are most common. In three, four and five dimensions, only a limited number of regular figures exist. In fact, in five and more dimensions, only three regular figures exist. They are, the regular simplex, the crosspolytope or p-dimensional octahedron, and the hypercube or p-dimensional cube. (See for example, Regular Polytopes, Coxeter, H.S.M. (1963), The MacMillian Company, New York, p. 291-296.)

Box and Hunter [6] point out that the crosspolytope and hypercube can always be combined to form a second-degree design. This combination of the crosspolytope and hypercube is an obvious extension of our previously mentioned two-level factorial design, and therefore, the combination will now be considered.

7.3 The central composite design.

In Chapters 5 and 6, the discussion on first-degree designs consisted of simplex designs as well as two-level factorial designs. If we confine our discussion at the

moment to the two-level factorial designs, we notice that these designs permit the estimation only of the coefficients α_0, α_i, α_{ij} $(1\leq i\leq p,\ 1\leq i<j\leq p)$ in the model

$$y = \alpha_0 + \sum_{i=1}^{p} \alpha_i w_i + \sum_{i=1}^{p} \alpha_{ii} w_i^2 + \sum\sum_{i<j}^{p} \alpha_{ij} w_i w_j + \epsilon. \quad (7.3.1)$$

With the addition of center point replicates however, it is shown in Appendix I that the estimate of $\alpha_{11}+\alpha_{22}+\ldots+\alpha_{pp}$ can be obtained but the experimenter is still without information on the individual α_{ii} $(1\leq i\leq p)$. It was with this latter object in mind (i.e., estimating the individual α_{ii}) that central composite designs were introduced by Box and Wilson in 1951.

In their original form, the central composite designs were formed by adding certain additional points to the two-level factorial designs. In fact, apart from the factorial portion of the design, these augmented designs (sometimes referred to as "star" designs) may be regarded as a one-factor-at-a-time experiment. This is because the experimenter starts at the center of the region of interest ($w_{ui}=0$, $1\leq i\leq p$) and then tests each factor or variable in turn at a high level ($w_{ui}=\omega$) and at a low level ($w_{ui}=-\omega$), while all the other variables are held at their central value ($w_{ui}=0$).

For example, with p variables, the central composite design which consists of the hypercube, the crosspolytope and n_0 center point replicates will have a design matrix of

the form

$$\underset{\sim}{D}_w = \begin{bmatrix} \pm c & \pm c & \cdots & \pm c \\ \pm c & \pm c & \cdots & \pm c \\ \vdots & \vdots & & \vdots \\ \pm c & \pm c & \cdots & \pm c \\ -\omega & 0 & \cdots & 0 \\ \omega & 0 & \cdots & 0 \\ 0 & -\omega & & 0 \\ 0 & \omega & & 0 \\ 0 & 0 & & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & & -\omega \\ 0 & 0 & & \omega \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \underset{\sim}{D}_1 \\ \\ \\ \\ \\ \underset{\sim}{D}_2 \\ \\ \\ \\ \\ \underset{\sim}{D}_3 \end{matrix} \quad . \qquad (7.3.2)$$

In $\underset{\sim}{D}_w$, which we horizontally partitioned into the matrices $\underset{\sim}{D}_1$, $\underset{\sim}{D}_2$ and $\underset{\sim}{D}_3$, the elements of the M x p matrix $\underset{\sim}{D}_1$ consist of the scalar product of the radius multiplier $\underline{c}$ and the plus and minus ones in the design matrix of a two-level factorial design. The matrix $\underset{\sim}{D}_2$ is a 2p x p matrix whose elements are the axial or star points. The n_0 x p matrix $\underset{\sim}{D}_3$ contains n_0 center point replicates. Hence, the total number of experiments to be performed with the rotatable central composite design is $N=M+2p+n_0$.

The factorial portion $\underset{\sim}{D}_1$ of the matrix $\underset{\sim}{D}_w$ will contain one or more first-degree orthogonal designs. That is,

the matrix $\underset{\sim}{D}_1$ will be part or all of the design matrix $\underset{\sim}{D}_w$ considered in Chapters 5 and 6. The n_0 center point replicates in $\underset{\sim}{D}_3$ consist of center point replicates performed with the factorial portion $\underset{\sim}{D}_1$ plus center point replicates run while performing the axial group $\underset{\sim}{D}_2$. Since the axial group is also a first-degree orthogonal design, then, as previously mentioned in Chapter 3, the central composite design is simply two or more first-degree designs augmented with an additional group of center point replicates.

If we refer back to (7.1.4), and with the design matrix defined by (7.3.2), we see that the matrix $\underset{\sim}{W}$, for general q, is now of the form

$$\underset{\sim}{W} = \begin{bmatrix} \underset{\sim}{D}_1 & c^2 \underset{\sim}{j}_M \underset{\sim}{j}_p' & \pm\sqrt{2c^2} & z_1, z_2, \ldots, z_{q-1} \\ & \begin{matrix} -\omega & 0 & \ldots & 0 \\ \omega & 0 & \ldots & 0 \\ 0 & -\omega & & 0 \\ 0 & \omega & & 0 \end{matrix} & & \\ \underset{\sim}{j}_N \quad \underset{\sim}{D}_2 & \begin{matrix} \vdots & \vdots & & \vdots \\ 0 & 0 & & -\omega \\ 0 & 0 & & \omega \end{matrix} & \underset{\sim}{0} & \\ & \begin{matrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \end{matrix} & & \\ \underset{\sim}{D}_3 & \begin{matrix} \vdots & \vdots & & \vdots \\ 0 & 0 & & 0 \end{matrix} & \underset{\sim}{0} & \end{bmatrix} \quad . (7.3.3)$$

The general form of the moment matrix using the central composite design is defined by (7.1.5) but now the quantities in (7.1.6) are

$$\lambda_2 = \frac{1}{N}[Mc^2 + 2\omega^2]$$

$$\lambda_4 = \frac{1}{N}[Mc^4]$$

$$G = \frac{1}{N}[Mc^4 + 2\omega^4] \qquad (7.3.4)$$

$$P_i = \frac{1}{N}\sum_{u=1}^{N} z_{q-1,u} w_{ui}^2 \qquad (1 \le i \le p)$$

$$N\underset{\sim}{Q} = \text{diag}\left(\sum_{u=1}^{N} z_{iu}^2\right) \qquad (1 \le i \le q-2)$$

$$d = \frac{1}{N}\sum_{u=1}^{N} z_{q-1,u}^2 \quad .$$

With regard to the quantities in (7.3.4) above, since the design points must lie within or on the boundary of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit inside the factor space, and if we put

$$\omega^2 = \rho_{ccd}^2$$

and

$$pc^2 \le \omega^2 \le \rho^{*2}, \qquad (7.3.5)$$

then,

$$c \le \frac{\omega}{\sqrt{p}} = \frac{\rho_{ccd}}{\sqrt{p}} \le \frac{\rho^*}{\sqrt{p}}$$

where ρ_{ccd} is the radius of the central composite design and ρ^* is the radius of the largest sphere. (Note that ρ^* could be the radius of the unit sphere. This choice is left to the experimenter as in Chapters 5 and 6.)

Probably the two most appealing characteristics of the central composite designs then are that they are not only easy to construct by adding points to the two-level factorial designs, but also, they do not contain an excessive number of design points. In addition, these designs permit the experimentation to be performed in stages. For example, if after the initial first-degree design (two-level factorial) is completed and one desires to fit a second-degree model, then all that is needed to form the central composite design is a few additional axial points $w_{ui} = \pm\omega$ $(1 \le i \le p)$ and some center point replicates. Here a stage might correspond to a period of time.

In our definition of the design matrix (7.3.2), we stated that the M x p matrix $\underset{\sim}{D}_1$ contained one or more first-degree orthogonal designs. Let us assume that the rows of $\underset{\sim}{D}_1$ comprise enough compartments, each containing a first-degree orthogonal design, so that with the z_{iu} $(1 \le i \le q-2, 1 \le u \le M)$ defined as in Section 3.3, $z_{i1}+z_{i2}+\ldots+z_{iM}=0$ $(1 \le i \le q-2)$. Then from the definition of $z_{q-1,u}$ in Section 3.3, we see that

$$z_{q-1,1} + z_{q-1,2} + \ldots + z_{q-1,M} = M.$$

On the other hand, if $\underset{\sim}{D}_1$ contains the complete factorial (i.e., all of the rows of $\underset{\sim}{D}_w$ in Chapters 5 and 6), then

$$z_{q-1,1} + z_{q-1,2} + \ldots + z_{q-1,M} = 0.$$

Let us assume that the design points in the rows of

the matrix $\underset{\sim}{D}_1$ as well as some center point replicates have been performed and refer to these experiments as the first stage of the central composite design. We are now ready to run the second stage $\underset{\sim}{D}_2$ as well as some additional center point replicates and at the same time, we shall want to measure the variation (if any exists) between the two stages of experimentation. Since we have discussed the use of orthogonal contrasts when fitting a first-degree model and since the elements in the columns of the matrix $\underset{\sim}{W}$ corresponding to the q-1 contrasts are already determined by the previous experimentation, we shall continue in the next section with the orthogonal contrast approach. It will be shown in the next section that orthogonality of the contrast column, which we use to measure the variation between the stages of experimentation, with the other columns in the matrix $\underset{\sim}{W}$ will influence the settings of the levels of the axial points in the rows of the matrix $\underset{\sim}{D}_2$.

7.4 The use of orthogonal contrasts to measure the variation between stages of experimentation.

Although we have referred to the completion of the factorial portion $\underset{\sim}{D}_1$ of $\underset{\sim}{D}_w$ in (7.3.2) as the first stage of experimentation, we can just as easily consider it the result of at least one stage. That is to say, up to now each stage of the experimentation process has consisted of some number of rows of the matrix $\underset{\sim}{D}_1$ as well as $n_{0j} \geq 0$

center point replicates. The running of the design points in the matrix $\underset{\sim}{D}_2$ as well as $n_{0\omega}$ center point replicates will then be the last stage of experimentation associated with the central composite design.

To obtain a measure of the variation that might exist between the stages of experimentation, we shall use one contrast containing N non-zero coefficients. In fact, since we are going to use just one contrast, then the coefficients z_u $(1 \leq u \leq N)$ in this contrast will not require a subscript i as in Section 3.3. To simplify the mathematics, let us replace the non-unit coefficients (if any exist) of the $q\text{-}1^{st}$ contrast in the matrix $\underset{\sim}{W}$ with the unit value, that is, $z_{q-1,u}=z_u=1$ $(1 \leq u \leq M)$. In the chapters on first-degree designs, the matrix $\underset{\sim}{D}_w$ was divided horizontally into q compartments and N_j $(1 \leq j \leq q)$ was the number of rows in the j^{th} compartment. We shall specify now that N_j is not only the number of rows of the j^{th} compartment of $\underset{\sim}{D}_1$, but N_j also includes n_{oj} center point replicates which combined with the rows of the j^{th} compartment of $\underset{\sim}{D}_1$ comprise and will be called the j^{th} group of design points. If we state that the matrix $\underset{\sim}{D}_1$ contains $\hat{q} \leq q$ compartments, then the axial points and $n_{0\omega}$ additional center point replicates will be the $(\hat{q}+1)^{st}$ group of experiments.

To see what values we impose on the elements ω in the matrix $\underset{\sim}{D}_2$ by making the N x 1 contrast column orthogonal to the other columns of the matrix $\underset{\sim}{W}$, let

$$T_j = \sum_{u=N_{j-1}+1}^{N_j} z_u = c_j(N_j) \qquad (1 \le j \le \hat{q}+1), \qquad (7.4.1)$$

and

$$Z_j = \sum_{u=N_{j-1}+1}^{N_j} z_u w_{ui}^2 = c_j \sum_{u=N_{j-1}+1}^{N_j} w_{ui}^2 \qquad (1 \le j \le \hat{q}+1, 1 \le i \le p), \qquad (7.4.2)$$

where z_u is the u^{th} coefficient in the contrast and c_j is a scalar quantity denoting the sign and magnitude of the coefficients of the contrast corresponding to the j^{th} group. Since each compartment of $\underset{\sim}{D}_1$ contains a first-degree orthogonal design and $z_u=1$ $(1 \le u \le N_j, 1 \le j \le \hat{q})$, then

$$\sum_{u=N_{j-1}+1}^{N_j} z_u w_{ui} w_{uj'} = \sum_{u=N_{j-1}+1}^{N_j} w_{ui} w_{uj'} \qquad (i \ne j')$$

$$= 0 \, . \qquad (7.4.3)$$

That is, the contrast column is orthogonal to the columns of the matrix $\underset{\sim}{W}$ corresponding to the crossproduct terms (see (7.3.3)).

Let us denote by $z_.^{(\omega)}$ the coefficients that fall in the rows corresponding to the rows of the matrix $\underset{\sim}{D}_2$ as well as the $n_{0\omega}$ center point replicates. Since we want the contrast column to be orthogonal to the columns corresponding to the pure quadratic terms $(w_{ui}^2; 1 \le i \le p)$ in the model (this will make $\underset{\sim}{P}=\underset{\sim}{0}$ in (7.1.5) and will simplify the inversion of the moment matrix as we shall see later), that is,

$$\sum_{u=1}^{N} z_u w_{ui}^2 = \sum_{j=1}^{\hat{q}+1} Z_j \qquad (1 \le i \le p)$$

$$= 0 , \qquad (7.4.4)$$

then

$$\sum_{u=N_{\hat{q}}+1}^{N_{\hat{q}+1}} z_{\cdot}^{(\omega)} w_{ui}^2 = - \sum_{j=1}^{\hat{q}} Z_j . \qquad (7.4.5)$$

The summation on the left-hand side of (7.4.5) is over the $z_u w_{ui}^2$ where u corresponds to the rows of $\underset{\sim}{D}_2$ plus $n_{0\omega}$ center point replicates (the axial group). By definition of z_u, that is, $T_1+T_2+\ldots+T_{\hat{q}+1}=0$, then

$$(2p+n_{0\omega}) z_{\cdot}^{(\omega)} = - \sum_{j=1}^{\hat{q}} \sum_{u=N_{j-1}+1}^{N_j} z_u$$

$$= - \sum_{j=1}^{\hat{q}} N_j z_{\cdot}^{(j)} , \qquad (7.4.6)$$

where $z_{\cdot}^{(j)}$ is the value of z_u for all N_j coefficients in the j^{th} group. But $z_{\cdot}^{(j)}=1$ for $1 \le j \le \hat{q}$, and therefore in order to make $z_1+z_2+\ldots+z_N=0$,

$$z_{\cdot}^{(\omega)} = - \frac{1}{(2p+n_{0\omega})} \sum_{j=1}^{\hat{q}} N_j$$

$$= - \frac{(N-N_\omega)}{N_\omega} , \qquad (7.4.7)$$

where $N_\omega = 2p+n_{0\omega}$.

To make the column of contrast coefficients orthogonal to the columns corresponding to the pure quadratic terms in the model, since $z_{\cdot}^{(\omega)}$ is constant from (7.4.7), then from (7.4.5) and (7.4.2),

$$\sum_{u=N_{\hat{q}}+1}^{N_{\hat{q}+1}} w_{ui}^2 = \left(\frac{N_\omega}{N-N_\omega}\right) \sum_{j=1}^{\hat{q}} Z_j \qquad (1 \le i \le p)$$

and

$$\sum_{u=N_{\hat{q}}+1}^{N_{\hat{q}+1}} w_{ui}^2 = \left(\frac{N_\omega}{N-N_\omega}\right) \sum_{j=1}^{\hat{q}} \sum_{u=N_{j-1}+1}^{N_j} w_{ui}^2. \qquad (7.4.8)$$

Therefore,

$$\frac{\sum_{u=N_{\hat{q}}+1}^{N_{\hat{q}+1}} w_{ui}^2}{\sum_{j=1}^{\hat{q}} \sum_{u=N_{j-1}+1}^{N_j} w_{ui}^2} = \frac{N_\omega}{N-N_\omega} \qquad (1 \le i \le p). \qquad (7.4.9)$$

Equation (7.4.9) shows that the ratio of the sum of the squares of the w_{ui} in the axial group to the sum of the squares of the w_{ui} in all the other groups must be directly proportional to the ratio of the corresponding number of design points that make-up the axial group and the sum of the design points in all of the other groups. Actually, equation (7.4.9) is used to determine the number of center point replicates $n_{0\omega}$ to be taken since the w_{ui}^2 $(1 \le u \le N, 1 \le i \le p)$

are usually fixed by some other criterion such as rotatability. We shall clarify this point with an example later in this section.

We have been talking about only one contrast containing N non-zero elements. Actually, the matrix $\underset{\sim}{W}$ may contain more than one contrast where the other contrasts all have zero value coefficients in the rows corresponding to the rows of the matrix $\underset{\sim}{D}_2$. (These contrasts formed the elements of the matrix $\underset{\sim}{Q}$ in (7.1.5).) In this case, a result which is identical to (7.4.9) can be derived involving the $\hat{q}$ groups in $\underset{\sim}{D}_1$ of $\underset{\sim}{D}_w$ only. For example, let the ℓ^{th} $(1\leq\ell\leq q-2)$ contrast column be orthogonal to the columns corresponding to the pure quadratic terms in the matrix $\underset{\sim}{W}$. Since the q-2 contrasts are already orthogonal to the "mean" column containing N ones, then they are also orthogonal to the column corresponding to the contrast L_{q-1} just derived. From (3.2.7), we know that

$$z_{\ell u}{}^{(\ell+1)} = -\frac{1}{N_{\ell+1}} \sum_{j=1}^{\ell} N_j \qquad (1\leq\ell\leq q-2) \qquad (7.4.10)$$

where $z_{\cdot}{}^{(j)}=1$ $(1\leq j\leq\ell)$ by definition in Section 3.3. Now, since the ℓ^{th} contrast is orthogonal to the columns containing the pure quadratic elements where N_j includes the center point replicates associated with the j^{th} group $(1\leq j\leq\hat{q})$, then

$$\sum_{u=N_\ell+1}^{N_{\ell+1}} z_{\ell u}{}^{(\ell+1)} w_{ui}^2 = -\sum_{j=1}^{\ell} \sum_{u=N_{j-1}+1}^{N_j} w_{ui}^2 \quad (1\leq i\leq p),$$

and therefore,

$$\frac{\sum_{u=N_{\ell}+1}^{N_{\ell+1}} w_{ui}^2}{\sum_{j=1}^{\ell}\sum_{u=N_{j-1}+1}^{N_j} w_{ui}^2} = \frac{N_{\ell+1}}{\sum_{j=1}^{\ell} N_j} . \qquad (7.4.11)$$

We now have in the matrix $\underset{\sim}{W}$, q-2 contrast columns whose elements are defined by (3.3.6), (3.3.6A) and (3.3.7), and one contrast column whose elements are defined by $z_{\cdot}^{(j)}=1$ where $j=1,2,\ldots,\hat{q}$ and equation (7.4.7).

Looking at the form of the design matrix (7.3.4) for a central composite design, we see that

$$\sum_{u}^{N_\omega} w_{ui}^2 = 2\omega^2 \qquad (1\le i\le p)$$

where the summation is over the elements in the axial group only. If we put this result in equation (7.4.9), the value of ω such that the contrast column z_{q-1} is orthogonal to the pure quadratic columns, making $\underset{\sim}{P}=\underset{\sim}{0}$ in (7.1.5), can be calculated. Actually, once the sizes of the N_j $(1\le j\le\hat{q})$ are determined, we could take the value of ω so that the design is rotatable. This is found by putting $G=3\lambda_4$ in (7.3.4), fixing the value of ω for rotatability at $\omega=\sqrt[4]{Mc}$. With this value of ω substituted in (7.4.9), the number $n_{0\omega}$ of center point replicates associated with the axial group is determined.

Equation (7.4.7) is then used to obtain the value of $z_{.}^{(\omega)}$ so that the contrast column is also orthogonal to the "mean" column in $\underset{\sim}{W}$.

Let us look at an example where p=2 and the matrix $\underset{\sim}{D}_1$ is of the form

$$\underset{\sim}{D}_1 = c \begin{bmatrix} -1 & -1 \\ -1 & 1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

In an attempt to obtain an estimate of error and some measure of the lack of fit of the first-degree model, two center point replicates were taken so that $N_1=6$. Now the value of ω in the matrix $\underset{\sim}{D}_2$ for rotatability is $\sqrt[4]{Mc} = \sqrt{2c}$. Putting this value of ω in (7.4.9), we see that the number of center point replicates $n_{0\omega}$ is

$$\frac{4 + n_{0\omega}}{6} = \frac{2(2c^2)}{4c^2}$$

$$n_{0\omega} = 2.$$

The value of $z_{.}^{(\omega)}$ so that the contrast column is orthogonal to the mean column in $\underset{\sim}{W}$ is, from (7.4.7),

$$z_{.}^{(\omega)} = -\frac{(12 - 6)}{6}$$

$$= -1,$$

and therefore the matrix $\underset{\sim}{W}$ is

$$\underset{\sim}{W} = \begin{bmatrix} 1 & -c & -c & c^2 & c^2 & c^2 & 1 \\ 1 & -c & c & c^2 & c^2 & -c^2 & 1 \\ 1 & c & -c & c^2 & c^2 & -c^2 & 1 \\ 1 & c & c & c^2 & c^2 & c^2 & 1 \\ 1 & -\sqrt{2}c & 0 & 2c^2 & 0 & 0 & -1 \\ 1 & \sqrt{2}c & 0 & 2c^2 & 0 & 0 & -1 \\ 1 & 0 & -\sqrt{2}c & 0 & 2c^2 & 0 & -1 \\ 1 & 0 & \sqrt{2}c & 0 & 2c^2 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix}.$$

Table II. on the next page shows the value of ω for rotatability as well as the value of ω for orthogonality of the contrast column for different p, $n_{0\omega}$, n_{0c} and the number of groups $\hat{q}$ in $\underset{\sim}{D}_1$. The quantity n_{0c} is the sum of all the center point replicates taken while performing the experiments in $\underset{\sim}{D}_1$. From the table, we see that when p=3,p=6 ($\frac{1}{2}$ rep), there is no way to satisfy rotatability and orthogonality simultaneously. In this case, a choice has to be made as to which criterion to sacrifice and we shall now consider the sacrificing of each criterion.

The first case is that of sacrificing the rotatability criterion for the ease in calculating the precision matrix (the inverse of the moment matrix). The orthogonality criterion is also necessary in order to get estimates (i.e., separate estimates) of all the parameters in the model. Although the variance contours will not be exact spheres,

Table II.

The Value of ω for Rotatability and Orthogonality

When Using the Central Composite Design.

p =	2	3	4	5 ($\frac{1}{2}$ rep.)	6 ($\frac{1}{4}$ rep.)	6 ($\frac{1}{2}$ rep.)	6
$\hat{q}$ =	1	2	2	2	2	2	8
$\sum_{j=1}^{\hat{q}} N_j$ =	6	12	20	22	26	40	72
n_{0c} =	2	4	4	6	10	8	8
$n_{0\omega}$ =	2	2	2	1	1	2	6
$z_\cdot^{(\omega)}$ =	-1	$-\frac{3}{2}$	-2	-2	-2	$\frac{-20}{7}$	-4
$\omega_{rot.}$ =	$\sqrt{2}c$	$\sqrt[4]{8}c$	2c	2c	2c	2.378c	$2\sqrt{2}c$
$\omega_{orth.}$ =	$\sqrt{2}c$	$\frac{\sqrt{8}}{\sqrt{3}}c$	2c	2c	2c	$\sqrt{5.6}c$	$2\sqrt{2}c$

the difference is slight (as evidenced by the small deviation of ω in the table). Thus the estimates of the constant parameter as well as the pure quadratic parameters, as noted by Box and Hunter, will be affected slightly as the design is rotated.

The second case is the sacrificing of the orthogonality criterion for the maintaining of spherical variance contours. The design is rotatable but the non-orthogonality property of the design is maintained in all rotations of the design and therefore, the estimates of the parameters α_0 and α_{ii} $(1 \leq i \leq p)$ have to be corrected for the contrast effect.

Since these arguments may be clarified by looking at the respective precision matrices, we shall take the space to show these precision matrices. The form of each precision matrix will, as we shall show, depend on the criterion used.

The form of the precision matrix when ω for orthogonality is chosen.

Suppose for p=3, p=6 ($\frac{1}{2}$ replicate), the value of ω is chosen so that the contrast column is orthogonal to the columns of the matrix $\underset{\sim}{W}$ corresponding to the pure quadratic terms in the model. Then the moment matrix is of the form (7.1.5) but where $\underset{\sim}{P}=\underset{\sim}{0}$. Since we know that the precision matrix is the inverse of the moment matrix, then to calculate $N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}$, we put $N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}[\frac{1}{N}(\underset{\sim}{W}'\underset{\sim}{W})]=\underset{\sim}{I}$. The form of the precision matrix is

$$N(\underset{\sim}{W}'\underset{\sim}{W})^{-1} = \begin{bmatrix} e & \underset{\sim}{0}' & f\underset{\sim}{j}_p' & \underset{\sim}{0}' & \underset{\sim}{0}' & 0 \\ \underset{\sim}{0} & \frac{1}{\lambda_2}\underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ & & \begin{matrix} g & h & h & \cdots & h \\ h & g & h & \cdots & h \\ h & h & g & & h \\ \vdots & \vdots & & \ddots & \vdots \\ h & h & & & h \\ h & h & \cdots & h & g \end{matrix} & & & \\ f\underset{\sim}{j}_p & \underset{\sim}{0} & & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \frac{1}{2\lambda_4}\underset{\sim}{I}_{\binom{p}{2}} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0}^{-1} & \underset{\sim}{0} \\ 0 & \underset{\sim}{0}' & \underset{\sim}{0}' & \underset{\sim}{0}' & \underset{\sim}{0}' & \frac{1}{d} \end{bmatrix}, \quad (7.4.12)$$

where the quantities in (7.4.12) are

$$\begin{aligned} e &= \frac{G + \lambda_4(p-1)}{U} \\ f &= \frac{-\lambda_2}{U} \\ g &= \frac{G + \lambda_4(p-2) - \lambda_2^2(p-1)}{(G - \lambda_4)\,U} \\ h &= \frac{\lambda_2^2 - \lambda_4}{(G - \lambda_4)U} \end{aligned} \quad (7.4.13)$$

and $U = G + \lambda_4(p-1) - p\lambda_2^2$.

Now when $p=2,4,5(\frac{1}{2}$ rep.$),6,6(\frac{1}{4}$ rep.$)$, the value of ω is the same for both criteria and therefore $G=3\lambda_4$ in (7.1.5) and (7.4.13). Thus the quantities in (7.4.13) are simplified to

$$e = \frac{\lambda_4(p+2)}{U'}$$

$$f = \frac{-\lambda_2}{U'} \qquad (7.4.14)$$

$$g = \frac{1}{2\lambda_4} + h$$

$$h = \frac{\lambda_2^2 - \lambda_4}{2\lambda_4 U'}$$

where $U' = \lambda_4(p+2) - p\lambda_2^2$. Note that when the design consists of just one single set of points which lie on the boundary of a p-dimensional sphere, then from (7.2.7) where

$$\lambda_4 = \frac{p\lambda_2^2}{p+2},$$

the quantity $U'=0$ and therefore e,f and h in (7.4.14) are undetermined.

<u>The form of the precision matrix when the value of ω for rotatability and not orthogonality is used</u>.

If one chooses the value of ω for rotatability when $p=3$, $p=6$ ($\frac{1}{2}$ rep.), then the moment matrix will be of the form (7.1.5) where $G=3\lambda_4$. The form of the precision matrix then is

$$N(\underset{\sim}{W}'\underset{\sim}{W})^{-1} = \begin{bmatrix} e & \underset{\sim}{0}' & f\underset{\sim}{j}_p' & \underset{\sim}{0}' & \underset{\sim}{0}' & i \\ \underset{\sim}{0} & \frac{1}{\lambda_2}\underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} \\ f\underset{\sim}{j}_p & \underset{\sim}{0} & \begin{matrix} g & h & h & \dots & h \\ h & g & h & \dots & h \\ h & h & g & & h \\ \vdots & \vdots & & & \vdots \\ h & h & & & h \\ h & h & \dots & h & g \end{matrix} & \underset{\sim}{0} & \underset{\sim}{0} & \tilde{\underset{\sim}{P}} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \frac{1}{2\lambda_4}\underset{\sim}{I}_{\binom{p}{2}} & \underset{\sim}{0} & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{Q}^{-1} & \underset{\sim}{0} \\ i & \underset{\sim}{0}' & \tilde{\underset{\sim}{P}}' & \underset{\sim}{0}' & \underset{\sim}{0} & \ell \end{bmatrix}, \quad (7.4.15)$$

where the quantities in (7.4.15) are

$$\begin{aligned} e &= 1 + \frac{p\lambda_2^2}{Y} \\ f &= \frac{-\lambda_2}{Y} \\ g &= \frac{(1-p)\lambda_2^2+(1+p)\lambda_4+(1-p)\frac{P}{d}^2}{2\lambda_4 Y} \\ h &= \frac{\lambda_2^2-\lambda_4+\frac{P}{d}^2}{2\lambda_4 Y} \\ i &= \frac{pP\lambda_2}{dY} \\ \tilde{P} &= \frac{-P}{dY} \\ \ell &= \frac{p(\lambda_4-\lambda_2^2)+2\lambda_4}{dY} \end{aligned} \quad (7.4.16)$$

and $Y = -p\lambda_2^2+(p+2)\lambda_4-p\frac{P}{d}^2$. It is easy to see that the matrix (7.4.12) is a much less complicated matrix than that of

(7.4.15). Based on this argument then, one might prefer to sacrifice the rotatability criterion when p=3 and p=6 ($\frac{1}{2}$ rep.).

We have been discussing the experimental program associated with the central composite design as well as the use of one orthogonal contrast for obtaining a measure of the between stages variation if the experimentation is performed in multiple stages. We have primarily been concerned with a stage as being a particular period of time in which a group of experiments is performed. As we said before however, a stage is any condition which is considered to be homogeneous during the running of one or more groups of experiments, for example, a period of time, a piece of equipment, etc.. We also mentioned before that the formulas for the estimates $\hat{\underset{\sim}{\alpha}}$ of $\underset{\sim}{\alpha}$ in (7.1.3) are derived in Appendix H. In most cases, the estimates $\hat{\underset{\sim}{\alpha}}$ can be found with the use of a simple desk calculator. This is verified by the example shown at the end of Chapter 8.

We have failed to mention that although the second-degree model and designs are formed by augmenting the first-degree model and designs in the system of the design variables, the experimental settings of the mixture components will be calculated exactly as was shown with the first-degree designs. For example, if the matrix $\underset{\sim}{W}$ is of the form (7.1.4) where the elements of $\underset{\sim}{D}_w$ are denoted by w_{uj} ($1 \leq u \leq N$; $2 \leq j \leq k-q+1$), and the elements of the matrix $\underset{\sim}{T}_1$ are denoted by t_{ij} ($1 \leq i \leq k$, $1 \leq j \leq k-q$), then from equation (3.2.15),

$$x_{uj} = \left(\sum_{\ell=2}^{k-q+1} w_{u\ell} t_{j,\ell-1} \right) h_j + x_{0j} \qquad (7.4.17)$$

is the level of the j^{th} component to be run in the u^{th} experiment.

Thompson and Myers [22] derive optimal central composite designs using the average mean square error of $\hat{y}$ where the average variance is considered only. In other words, in assuming that most of the variation in the predicted response using a second-degree model is sampling variation and the bias is negligible, Thompson and Myers state that designs with both the axial (ω) and factorial points placed on the boundary of the unit sphere provide the smallest value of J=V for different values of ω, c, n_0 and p.

Box and Draper [4] also discuss the minimization of the average variance of $\hat{y}$ (assuming the bias of $\hat{y}$ is negligible) for composite designs. They show that to minimize the average variance criterion, the largest design possible (even with points outside the region of interest) should be used. In addition, they suggest adding center point replicates in order to obtain a better estimate of the error variance. Since we are concerned with achieving orthogonality of the contrast column in the $\underset{\sim}{W}$ matrix as well as rotatability, caution must be taken when using the tables tabulated by Thompson and Myers, as well as Box and Draper when seeking

optimal second-degree designs.

7.5 Simplex-sum designs.

In the last two sections, we discussed the central composite design with the assumption that for fitting a first-degree model, a two-level factorial design is used. It was stated in Chapter 5 however that when p=2 and the double-simplex design is used, then a simple addition of center point replicates would result in the formation of the second-degree hexagonal design. It is interesting to note that this particular design could also be derived by adding the vectors which connect the centroid with the vertices of the 2-dimensional simplex (equilateral triangle). This method was introduced by Box and Behnken [2] and the resulting designs are called simplex-sum designs. These designs however do not offer any advantages over the composite designs and in fact, require more experimental points in most cases and therefore, we shall discuss only the cases where p=2 and p=3 where p=k-q.

Let q=2 and consider the design points located only in the top half of the design matrix $\underset{\sim}{D}_w$ (Chapter 5). The number of experimental points is N=p+1 where p is the number of design variables. In Chapter 5, it was pointed out that in the design space of the w_i ($1 \leq i \leq p$), the design points lie at the vertices of an inscribed regular simplex. For example, if p=2, the design points lie at the vertices of an

equilateral triangle; if p=3, the points lie at the vertices of a regular tetrahedron.

When p=2, then the three points at the vertices of an equilateral triangle when joined to the origin of the design space (center of the triangle), define three vectors. If we add these vectors two at a time, a second equilateral triangle is obtained and by adding the original vectors three at a time, we obtain a center point. Therefore, the original set of points plus the derived points generate the hexagonal design.

When p=3, the four vectors from the center of the tetrahedron to the vertices of the tetrahedron when added in all possible ways two at a time generate six further vectors which pass through the midpoints of the edges of the tetrahedron. When the four vectors are added in all possible ways three at a time, they generate four vectors passing through the midpoints (centroids) of the faces of the tetrahedron. Finally, when the four vectors are all added together, a center point results. If the lengths of the derived vectors are suitably scaled, the resulting design coincides with the central composite design.

Since the method of constructing the simplex-sum designs is discussed in [2], we shall only briefly discuss the construction of a design matrix for the case p=3 simply to outline the method. If we let q=2 and denote the top half of $\underset{\sim}{D}_w$ in (5.1.2) by $\underset{\sim}{D}^{(1)}$, then

$$\underset{\sim}{D}^{(1)} = \begin{bmatrix} \underset{\sim}{w}_1' \\ \underset{\sim}{w}_2' \\ \underset{\sim}{w}_3' \\ \underset{\sim}{w}_4' \end{bmatrix} \tag{7.5.1}$$

where $\underset{\sim}{w}_i'$ is the i^{th} row vector of design points. If we take all combinations of two rows of $\underset{\sim}{D}^{(1)}$ at a time, we have

$$\underset{\sim}{D}^{(2)} = \begin{bmatrix} \underset{\sim}{w}_1' + \underset{\sim}{w}_2' \\ \underset{\sim}{w}_1' + \underset{\sim}{w}_3' \\ \underset{\sim}{w}_1' + \underset{\sim}{w}_4' \\ \underset{\sim}{w}_2' + \underset{\sim}{w}_3' \\ \underset{\sim}{w}_2' + \underset{\sim}{w}_4' \\ \underset{\sim}{w}_3' + \underset{\sim}{w}_4' \end{bmatrix} . \tag{7.5.2}$$

The matrix $\underset{\sim}{D}^{(2)}$ is referred to in [2] as the second subset of design points. Taking all combinations of sums of three rows of $\underset{\sim}{D}^{(1)}$ at a time, we have

$$\underset{\sim}{D}^{(3)} = \begin{bmatrix} \underset{\sim}{w}_1' + \underset{\sim}{w}_2' + \underset{\sim}{w}_3' \\ \underset{\sim}{w}_1' + \underset{\sim}{w}_2' + \underset{\sim}{w}_4' \\ \underset{\sim}{w}_1' + \underset{\sim}{w}_3' + \underset{\sim}{w}_4' \\ \underset{\sim}{w}_2' + \underset{\sim}{w}_3' + \underset{\sim}{w}_4' \end{bmatrix} \tag{7.5.3}$$

and $\underset{\sim}{D}^{(3)}$ is called the third subset of points. Geometrically these subsets are symmetrically oriented one to another in that the row vectors of $\underset{\sim}{D}^{(2)}$ bisect the edges of the simplex defined by $\underset{\sim}{D}^{(1)}$. The row vectors of $\underset{\sim}{D}^{(3)}$ pass through the centroids of the faces of the simplex defined by $\underset{\sim}{D}^{(1)}$.

To determine the radius of the points in each of the subsets of points, let

$$\underset{\sim}{D}_w = \begin{bmatrix} a_1 \underset{\sim}{D}^{(1)} \\ a_2 \underset{\sim}{D}^{(2)} \\ a_3 \underset{\sim}{D}^{(3)} \\ \underset{\sim}{n} \end{bmatrix} \qquad (7.5.4)$$

where the a_i $(1 \le i \le 3)$ can be called radius multipliers and $\underset{\sim}{n}$ is a vector of center point replicates the elements of which can be generated by $\underset{\sim}{w}_1' + \underset{\sim}{w}_2' + \underset{\sim}{w}_3' + \underset{\sim}{w}_4'$. Since the radius of the design points must be less than or equal to the radius of the largest sphere centered at $\underset{\sim}{w}=\underset{\sim}{0}$ that will fit in the factor space, then

$$a_i^2(\underset{\sim}{w}_j'\underset{\sim}{w}_j) \le \rho^{*2} \qquad (i=j,\ i\neq j). \qquad (7.5.5)$$

But from Chapter 5, we know that with the simplex design,

$$[\underset{\sim}{j}_{p+1},\ \underset{\sim}{D}^{(1)}]'[\underset{\sim}{j}_{p+1},\ \underset{\sim}{D}^{(1)}] = \begin{bmatrix} p+1 & \underset{\sim}{0}' \\ \underset{\sim}{0} & (p+1)\lambda_2 \underset{\sim}{I}_p \end{bmatrix}$$

and therefore,

$$\begin{aligned} \underset{\sim}{w}_i'\underset{\sim}{w}_j &= (p+1)\lambda_2 - 1 \quad i=j \\ &= -1 \qquad\qquad i\neq j. \end{aligned} \qquad (7.5.6)$$

If $\underset{\sim}{w}_{u(i)}'$ is the u^{th} row vector in the i^{th} subset, then the square of the length of $\underset{\sim}{w}_{u(i)}'$ is

$$\begin{aligned} \underset{\sim}{w}_{u(i)}'\underset{\sim}{w}_{u(i)}' &= i[(p+1)\lambda_2 - 1] + 2\binom{i}{2}[-1] \\ &= i[(p+1)\lambda_2 - i]\ . \end{aligned} \qquad (7.5.7)$$

However, since $a_i^2(\underset{\sim}{w}_j'\underset{\sim}{w}_j) \le \rho^{*2}$, then

$$a_i^2[(p+1)\lambda_2 - i]i \le \rho^{*2},$$

and

$$a_i \leq \frac{\rho^*}{[i(p+1)\lambda_2 - i^2]^{\frac{1}{2}}} . \qquad (7.5.8)$$

In [2], standard solutions are tabulated based on the convention that

$$\lambda_2 = \frac{1}{N}\sum_{u=1}^{N} w_{ui}^2 = 1, (1 \leq i \leq p).$$

For the standard solution in [2], the radius of the points in the i^{th} subset (i.e., a suitable set of radius multipliers such that the moments of the design fulfill the requirements for rotatability), is

$$r_i = \binom{(p+1) - 2}{i - 1}^{-\frac{1}{4}} [i(p+1-i)]^{\frac{1}{2}}. \qquad (7.5.9)$$

If ρ^* in (7.5.8) is replaced by r_i in (7.5.9), then a_i $(1 \leq i \leq 3)$, the radius multiplier for the i^{th} subset can be calculated.

VIII. SUMMARY, RECOMMENDATIONS AND INDUSTRIAL EXAMPLE.

8.1 Summary and recommendations.

In this paper, we have attempted to utilize known response surface methods (design configurations and the definition of a region of interest), when the experimentation involves mixture components which are classified categorically. That is to say, the mixture components belong to q distinct categories ($q \geq 2$) where each category contributes a fixed proportion to every mixture and therefore every category must be represented in the mixtures by one or more of its member components. This statement of the problem eliminates situations where a mixture may consist simply of one component or of some combination of components all of which belong to one particular category, for example, in the combinations of acids and glycols considered in Section 1.3 of Chapter 1, we could not consider a mixture consisting only of acid constituents.

By the restrictions placed on the proportions contributed by each of the categories (and therefore on the proportions contributed by the components in these categories), the factor space of the components loses its geometrical simplicity. That is to say, the factor space no longer assumes the form of a simplex configuration as in the usual (Scheffé) mixture problem, but rather the form of a convex polytope. The complexity of the polytopes in terms of the number of extremities (boundaries of the polytopes)

corresponding to two categories is shown in Section 2.3 of Chapter 2.

Unlike the statements of the problems of Scheffé [20] and Lambrakis [16] where the interest is extended over the entire factor space of the components, in our problem, the interest is defined to some smaller region within the factor space (convex polytope). The region of interest for the k mixture components is defined analytically to be ellipsoidal and of the form

$$\sum_{i=1}^{k} \left(\frac{x_i - x_{0i}}{h_i} \right)^2 \leq 1$$

where the x_{0i} and h_i $(1 \leq i \leq k)$ are chosen by the experimenter. The x_{0i} denotes the center of the interval of interest for the i^{th} $(1 \leq i \leq k)$ component and h_i is a constant which allows for the spread of the symmetric interval of interest for the i^{th} component. This definition of a region of interest enables us to isolate the experimentation to specific areas of interest and therefore ignore other areas of the factor space which are not of interest. The centroid $\underset{\sim}{x}_0$ of the ellipsoidal region is called the point of maximum interest and in some cases, this point may coincide with the centroid of the convex polytope, that is, when $x_{0i} = \frac{1}{k}$ $(1 \leq i \leq k)$.

To simplify the analysis in which we use least-squares theory as well as to center the problem around the point of maximum interest, a transformation of the mixture components

to intermediate variables is performed. The intermediate variables are defined by

$$v_j = \frac{x_j - x_{0j}}{h_j} \qquad (1 \le j \le k).$$

This transformation enables us not only to eliminate the physical units of the mixture components but also the problem is simplified in that a unit sphere (in the metric of the v_j's) which is centered at $v_j=0$ $(1 \le j \le k)$ is much easier to work with than is an ellipsoidal region especially when constructing designs.

The rank of the N x k matrix $\underset{\sim}{X}$ is k-q in the general linear model $\underset{\sim}{y} = \underset{\sim}{X}\ \underset{\sim}{\gamma} + \underset{\sim}{\varepsilon}$, where the elements of the matrix $\underset{\sim}{X}$ specify the levels of the k mixture components to be run in the N experiments. This less-than-full-rank property of the matrix $\underset{\sim}{X}$ is a consequence of the constraints placed on the components. Therefore, if

$$\underset{\sim}{v}_u' = (\underset{\sim}{x}_u - \underset{\sim}{x}_0)'\underset{\sim}{H}^{-1} \qquad (1 \le u \le N)$$

where $\underset{\sim}{x}_u' = (x_{u1}, x_{u2}, \ldots, x_{uk})$, $\underset{\sim}{x}_0' = (x_{01}, x_{02}, \ldots, x_{0k})$ and $\underset{\sim}{H} = \text{diag}(h_1, \ldots, h_k)$, is the u^{th} row vector of the N x k matrix $\underset{\sim}{V}$, then the rank of the matrix $\underset{\sim}{V}$ is also k-q. Since most of the work involving response surface methods has been developed under the assumption that the variables are independent and since the criterion used in this paper to optimize designs can be applied more easily using independent variables, a reparametrization of the model $\underset{\sim}{y} = \underset{\sim}{V}\ \underset{\sim}{\beta} + \underset{\sim}{\varepsilon}$ to a model

containing independent variables is performed. The reparametrization is performed by using an orthogonal matrix $\underset{\sim}{T}$ and the resulting model is $\underset{\sim}{y} = \tilde{\underset{\sim}{W}}\, \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}$, where $\underset{\sim}{\alpha}_3' = (\alpha_1, \ldots, \alpha_{k-q})$. The variables w_i $(1 \leq i \leq k-q)$ in this latter model are called the design variables.

By the method used to transform the mixture components x_i $(1 \leq i \leq k)$ to the v_i $(1 \leq i \leq k)$ as well as the transformation to the w_i $(1 \leq i \leq k-q)$, it is shown that once the estimates $\hat{\underset{\sim}{\alpha}}_3$ of the parameters in the model $\underset{\sim}{y} = \tilde{\underset{\sim}{W}}\, \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}$ are obtained, a simple inverse transformation enables us to represent the estimated response with either a model in the v_i's or a model in the x_i's. Thus the procedure is that we begin with dependent mixture components but work (construct designs, etc.) with independent design variables.

The N x (k-q) matrix $\tilde{\underset{\sim}{W}}$ in the model $\underset{\sim}{y} = \tilde{\underset{\sim}{W}}\, \underset{\sim}{\alpha}_3 + \underset{\sim}{\varepsilon}$ consists only of the levels of the design variables to be used in the N experiments. We show how q additional columns of non-zero elements can be added to the matrix $\tilde{\underset{\sim}{W}}$ enabling us to obtain information about the observations which would otherwise not be detected or measured. The addition of the q columns (i.e., the derivation of the elements of the q columns) is such that the estimates of the effects of the k-q design variables are unaffected by the addition of the q terms to the model. The addition of the q columns to the matrix $\tilde{\underset{\sim}{W}}$ is accomplished in Chapter 3 with the use of orthogonal contrasts of the observations and the theory is shown to be

analogous to the theory behind orthogonal blocking. The resulting matrix after the q columns are added is the N x k matrix $\underset{\sim}{W}$ and the general linear model is expressed as $\underset{\sim}{y} = \underset{\sim}{W}\,\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$. The matrix $\underset{\sim}{W}$ has rank k.

The criterion used for optimizing first-degree designs is the average mean square error of $\hat{y}$, introduced by Box and Draper [3]. This criterion involves the integration over a region of interest where for the most part the design points are placed. Since Dirichlet multiple integrals can be used for integrating over k-dimensional spherical regions, the integration over the region of interest is simplified. An extended region of interest defined as the largest sphere centered at the point of interest that will fit inside the factor space is derived. The radius (denoted by ρ^*) of this largest spherical region gives us an upper bound in terms of the distance that the design points can lie from the center of the design since the designs suggested (Chapters 5 and 7) for use are rotatable. (In so far as we have claimed that any matrix $\underset{\sim}{T}_1$ with k-q orthonormal columns such that the matrix $\underset{\sim}{T}$ is orthogonal can be used, we have in affect claimed that the design can itself be rotated. If any design points were outside this sphere, then upon rotation, they might fall outside the experimental region.)

In Chapter 5, the derivation of the optimal first-degree designs for the "all-variance" situation (where the bias of $\hat{y}$ is assumed to be zero), and the "all-bias"

situation (where the variance of $\hat{y}$ is assumed to be zero) are shown as are the settings of the levels of the design variables using a double-simplex design and a scaled two-level factorial design. The simplex and the two-level factorial designs are particularly convenient in that they possess the property that they can be augmented easily to second-degree designs. This property of the designs is verified in Chapter 7 when the Central Composite and Simplex-sum designs are discussed.

As we stated, rotatable designs are suggested for use. This is for two reasons. First, at the beginning of Chapter 5, it is explained that by using rotatable designs, we are able to use <u>any</u> k x (k-q) matrix $\underset{\sim}{T}_1$ with orthonormal columns so that the k x k matrix $\underset{\sim}{T}$ is orthogonal in the reparametrization of the model without affecting the properties (variance and bias) of the predictor $\hat{y}$. Secondly, the purpose of the experimental design is to explore a response surface and often to try to locate a maximum or minimum point when the orientation of the response surface with respect to the experimental region is unknown. For this reason, it would seem desirable to use an experimental design which provides uniform information, that is, constant Var $\hat{y}$ on each sphere which is centered at $\underset{\sim}{w}=\underset{\sim}{0}$. With rotatable designs, the Var $\hat{y}$ is a function only of the distance the predicted point lies from the center of the design regardless of the direction from the center of the design. This property is

especially desirable in initial experimentation.

Since most of the work in Chapters 3, 4 and 5 deals with the problem of just two categories of mixture components, in Chapter 6 the development is extended to three and more categories of components. The discussion is held to a minimum since most of the theory is an obvious extension of the theory involving two categories of mixture components.

In Chapter 7, the use of second-degree designs and second-degree models for measuring the curvature of the response surface is shown. The use of orthogonal contrasts for measuring variation that might exist between stages of experimentation is also extended to the second-degree situation.

Recommendations.

We have shown when using a first-degree model, how the estimated response can be expressed in the models of the mixture components, intermediate variables and the design variables. The reason for showing the different models is that applications of the mixture problem are diverse as are the personnel involved with the applications. Some may prefer to use the model in the mixture components while others prefer to use the model containing the design variables.

The feeling of the writer of this paper is that the latter model (the model in the design variables) is a more convenient model to work with not only for first degree but

for second-degree cases as well. This is because the response surface is defined with q fewer variables (q is the number of categories), and the estimates of the parameters in the model can be interpreted more easily. Therefore, when plotting contours of constant response, locating stationary points, etc., the labor is reduced since the terms corresponding to the contrast effects can be ignored during these operations.

Does this imply then that the development of the orthogonal contrasts of the observations can be ignored? On the contrary; dividing the program of experimentation into separate groups of experiments has been proven to be an effective method for studying response surfaces. It was also shown that with the use of orthogonal contrasts, variation could be removed from the analysis of the response variables when the experimentation is divided into groups of experiments.

The extension of the use of orthogonal contrasts when fitting a second-degree model could be omitted when the experimentation or execution of the second-degree design points is performed in one stage. This would eliminate conflicting settings of $\pm\omega$ when p=3 and p=6 ($\frac{1}{2}$ rep.), enabling the experimenter to concentrate solely on rotatability. However, when a possible source of variation exists and this variation can be measured when the experimentation is performed in separate groups of experiments, it would seem wise

to use the contrast terms in the model. The industrial problem in the next section exemplifies this point.

In the future, time will be spent on the inclusion of quantitative variables such as temperature, pressure, etc., in the mixture problem. The use of a different coordinate system for the design variables such as a complex coordinate system will also be researched in an attempt to preserve symmetry of the design configurations in the system of the mixture components. There are many additional areas, relating to the mixture problem, which should be investigated such as the use of designs other than rotatable designs and the generalization of the ellipsoidal region of interest so that the principal axes are not necessarily parallel to the axes of the mixture components.

8.2 An industrial example using a second-degree central composite design.

In the production of a polymeric solution, the mixture system involves combinations containing at least one acid with at least one glycol. For this particular example, the chemistry of the solution has been reduced to two acids x_1, x_2 and two glycols x_3, x_4. The purpose of the experimentation is to try to adequately describe the response surface with a prediction model so that future predictions involving these four components can be made. It is desired also to draw contours of the response surface so as to pic-

torially represent the response surface over the region of interest.

The only knowledge concerning the experimentation is that the combinations x_1x_4 and x_2x_4 produce unsatisfactory results. It is assumed that a quadratic model should adequately fit the response system and one stipulation in the experimental program is that the setting up as well as the running of the mixture combinations can be performed on only two pieces of chemical equipment. Thus, in addition to constructing a second-degree design, we want to group the experiments so as to remove any variation that exists between the two pieces of equipment.

It is decided to place the base point $\underset{\sim}{x}_0$ at the center of the factor space, that is, $\underset{\sim}{x}_0' = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. The components are of equal importance and thus given equal spread and $\underset{\sim}{H} = \text{diag}(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. Since we want to plot response contours over the region of interest and also want the largest possible region for this particular example, the design points are placed on the boundary of the largest spherical region. A central composite design is to be used and the radius multiplier c=1.0.

From Table II., the value of ω for rotatability as well as orthogonality when k-q=2 is

$$\omega = \sqrt{2}c$$

$$= 1.414.$$

Thus the matrix $\underset{\sim}{W}$ is

$$\underset{\sim}{W} = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 & 1.414 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1.414 & 1 \\ 1 & -1 & 1 & 1 & 1 & -1.414 & 1 \\ 1 & 1 & -1 & 1 & 1 & -1.414 & 1 \\ 1 & -1.414 & 0 & 2 & 0 & 0 & -1 \\ 1 & 1.414 & 0 & 2 & 0 & 0 & -1 \\ 1 & 0 & -1.414 & 0 & 2 & 0 & -1 \\ 1 & 0 & 1.414 & 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

and the matrix $\underset{\sim}{T}$ is

$$\underset{\sim}{T} = \begin{bmatrix} .707 & -.707 & 0 & 0 \\ .707 & .707 & 0 & 0 \\ 0 & 0 & -.707 & .707 \\ 0 & 0 & .707 & .707 \end{bmatrix} .$$

To get the corresponding settings of the mixture components, we know from (3.2.15),

$$x_{uj} = \left(\sum_{i=2}^{3} w_{ui} \, t_{ji} \right) h_j + x_{0j}$$

where t_{ji} is the element in the j^{th} row and i^{th} column of the matrix $\underset{\sim}{T}$. The levels of the mixture components and the values of the observations taken at these levels of the mixture components are

x_1	x_2	x_3	x_4	design point	response
.427	.073	.427	.073	(1)	9.3
.073	.427	.073	.427	(2)	8.2
.427	.073	.073	.427	(3)	6.1
.073	.427	.427	.073	(4)	10.4
.50	.00	.25	.25	(5)	6.4
.00	.50	.25	.25	(6)	8.9
.25	.25	.50	.00	(7)	11.9
.25	.25	.00	.50	(8)	7.3
.25	.25	.25	.25	(9)	7.7
.25	.25	.25	.25	(10)	7.9
.25	.25	.25	.25	(11)	7.8
.25	.25	.25	.25	(12)	7.8 .

If we refer to Appendix H, we see that the estimates $\hat{\alpha}$ in the second-degree model

$$y = \alpha_0 + \alpha_1 w_1 + \alpha_2 w_2 + \alpha_{11} w_1^2 + \alpha_{22} w_2^2 + \alpha_{12} w_1 w_2 + \alpha_{L_1} z_1 + \epsilon$$

are

$$\hat{\alpha}_0 = \frac{1}{4} \sum_{u=9}^{12} y_u = 7.80$$

$$\hat{\alpha}_1 = \frac{1}{8}(-y_1 + y_2 - y_3 + y_4) + \frac{\sqrt{2}}{8}(y_6 - y_5) = .8420$$

$$\hat{\alpha}_2 = \frac{1}{8}(-y_1 + y_2 + y_3 - y_4) + \frac{\sqrt{2}}{8}(y_8 - y_7) = -1.4882$$

$$\hat{\alpha}_{11} = \frac{1}{16} \sum_{u=1}^{4} y_u - \frac{1}{16} \sum_{u=5}^{8} y_u - \frac{1}{8} \sum_{u=9}^{12} y_u + \frac{1}{4}(y_5 + y_6) = -.1062$$

$$\hat{\alpha}_{22} = \frac{1}{16} \sum_{u=1}^{4} y_u - \frac{1}{16} \sum_{u=5}^{8} y_u - \frac{1}{8} \sum_{u=9}^{12} y_u + \frac{1}{4}(y_7 + y_8) = .8688$$

$$\hat{\alpha}_{12} = \frac{1}{4}(y_1 + y_2 - y_3 - y_4) = 0.25$$

$$\hat{\alpha}_{L_1} = \frac{1}{12}\left(\sum_{u=1}^{4} y_u + y_{11} + y_{12} - \sum_{u=5}^{10} y_u \right) = .0417$$

and therefore, the prediction model is

$$\hat{y} = 7.8 + .842 w_1 - 1.4882 w_2 - .1062 w_1^2 + .8688 w_2^2 + .25 w_1 w_2.$$

With this prediction model, the response contours were plotted and are shown in Figure 3 on the next page.

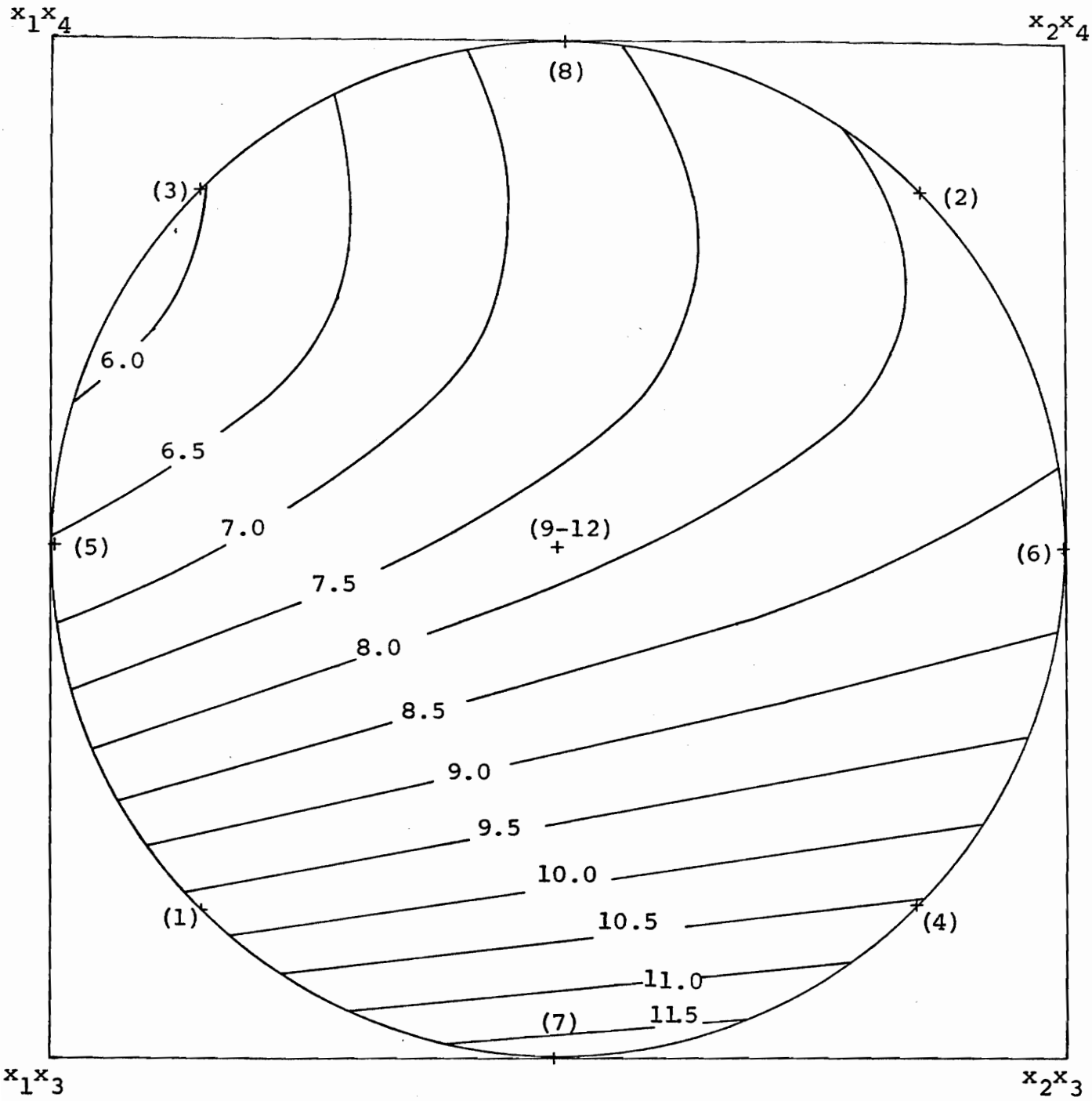

Figure 3.

Response Contours Over the Region of Interest.

BIBLIOGRAPHY

1. Box, G.E.P. (1952). Multi-Factor Designs of First Order. Biometrika, Vol. 39, 49-57.

2. Box, G.E.P. and Behnken, D.W. (1960). Simplex-Sum Designs: A Class of Second Order Rotatable Designs Derivable from Those of First Order. Annals of Math. Statistics, Vol. 31, 838-864.

3. Box, G.E.P. and Draper, N.R. (1959). A Basis for the Selection of a Response Surface Design. Jour. Amer. Statistical Assoc., Vol. 54, 622-654.

4. Box, G.E.P. and Draper, N.R. (1963). The Choice of a Second Order Rotatable Design. Biometrika, Vol. 50, 335-352.

5. Box, G.E.P. and Gardner, C.J. (1966). Constrained Designs. Technical Report No. 89, Dept. of Statistics, University of Wisconsin, Madison, Wisconsin.

6. Box, G.E.P. and Hunter, J.S. (1957). Multifactor Experimental Designs for Exploring Response Surfaces. Annals of Math. Statistics, Vol. 28, 195-241.

7. Box, G.E.P. and Wilson, K.B. (1951). On the Experimental Attainment of Optimum Conditions. Jour. of the Royal Stat. Soc., Series B, Vol. 13, 1-45.

8. Cochran, W.G. and Cox, G.M. (1964). Experimental Designs. Second Edition, John Wiley and Sons, Inc., New York, 62-67.

9. Davies, O.L. (1960). The Design and Analysis of Industrial Experiments. Oliver and Boyd, Hafner Publishing Company, New York, 344-352, 427-431.

10. Draper, N.R. and Lawrence, W.E. (1965). Mixture Designs for Three Factors. Jour. of the Royal Stat. Soc., Series B, Vol. 27, 450-454. Mixture Designs for Four Factors. Loc. cit. 473-478.

11. Finney, D.J. (1960). An Introduction to the Theory of Experimental Design. The University of Chicago Press, Chicago, Illinois, 15-17.

12. Gorman, J.W. and Hinman, J.E. (1962). Simplex Lattice Designs for Multicomponent Systems. Technometrics, Vol. 4, 463-487.

13. Graybill, F.A. (1961). An Introduction to Linear Statistical Models. Vol. 1, McGraw-Hill Book Company Inc., New York, 235-238.

14. Kendall, M.G. (1961). A Course in The Geometry of n Dimensions. No. 8, Griffins Statistical Monographs and Courses, Hafner Publishing Company, New York, 9-11.

15. Kurotori, I.S. (1966). Experiments with Mixtures of Components Having Lower Bounds. Industrial Quality Control, May, 592-596.

16. Lambrakis, D.P. (1968). Experiments with Mixtures: A Generalization of the Simplex-Lattice Design. Jour. of the Royal Stat. Soc., Series B, Vol. 30, 123-136.

17. McLean, R.A. and Anderson, V.L. (1966). Extreme Vertices Design of Mixture Experiments. Technometrics, Vol. 8, 447-454.

18. Plackett, R.L. (1960). Principles of Regression Analysis. Oxford University Press, Amen House, London E.C. 4, 33-44, 107-111.

19. Plackett, R.L. (1950). Some Theorems in Least Squares. Biometrika, Vol. 37, 149-157.

20. Scheffé, H. (1958). Experiments with Mixtures. Jour. of the Royal Stat. Soc., Series B, Vol. 20, 344-360.

21. Sommerville, D.M.Y. (1958). An Introduction to the Geometry of N Dimensions. Dover Publications, Inc., New York, 7-10, 96-106.

22. Thompson, W.O. and Myers, R.H. (1968). Response Surface Designs for Experiments with Mixtures. Technometrics, Vol. 10, 739-756.

APPENDIX A.

GLOSSARY OF TERMS.

This glossary is divided into two parts. Part I covers the material in Chapters 1 and 2 where in these chapters the problem of categorizing the mixture components is introduced and the literature concerning mixture experimentation is reviewed. Part II covers the material in Chapters 3 to 8 where the solution to the categorized mixture problem is discussed. The reason for dividing the glossary into two parts is that the notation overlaps in some cases. That is, the symbols used in Part II in some cases are the same as in Part I but the definitions are changed. The reason for changing the definitions is that in Part I, the notation in the literature review is kept as in the literature reviewed and in Chapters 3 to 8, the notation is used in a different sense because of the lack of new symbols which can be used. In the cases where the definition of a symbol used in Part II is the same as the definition used in Part I, the statement Part I is entered only in Part II under the heading of Definition.

Part I: Chapters 1 and 2.

	Symbol	Term	Definition
1.	β_i, γ_i	unknown parameters	These parameters are used in the polynomial representation of the unknown response surface.

	Symbol	Term	Definition
2.	$b_i, c_i, d_i, \hat{\beta}_i$	estimates	Parameters and parameter estimates used in the fitted models for estimating the unknown response surface.
3.	c_i	proportion	The proportion contributed by the i^{th} $(1 \leq i \leq n)$ major component in a mixture. Lambrakis [16].
4.	$\underset{\sim}{D}_x$	design matrix	A two-level factorial design matrix in k variables. It has N rows. Box and Gardner [5].
5.	$\underset{\sim}{D}_z$	design matrix	The N x k matrix of constrained variables. Box and Gardner [5].
6.	ε_u	error	The random error occuring in the u^{th} observation as a result of observing the true response.
7.	$\eta(\cdot)$	response	Unknown response which defines a response surface over some region.
8.	$f(\cdot)$	function	Polynomial expression used for estimating the true form of the response surface.
9.	h	scalar quantity	Upper bound for the proportion contributed by the first component in a mixture. Scheffé [20].

	Symbol	Term	Definition
10.	k	scalar quantity	The total number of mixture components forming the mixtures.
11.	K		The proportion or fraction of the total mixture contributed by the components in category 2.
12.	ℓ_i, ℓ_i'		Lower and upper bounds respectively for the i^{th} mixture component. McLean and Anderson [17].
13.	n_j		The number of components in the j^{th} $(1 \leq j \leq q)$ category. When $q=2$, $n_1=n$.
14.	m_i		Degree of polynomial used to represent the response over the i^{th} $(1 \leq i \leq n)$ simplex.
15.	μ		The number of mixture components set equal to zero for determining an extremity of dimensionality $k-q-\mu$ of the polytope.
16.	$\tilde{N}$, $\tilde{N}_i$		The number of boundaries of dimensionality $k-q-\mu$ and i respectively of the convex polytope.
17.	N		The number of experiments (design points) used for estimating the parameters in the model.

	Symbol	Term	Definition
18.	q	scalar quantity	The number of categories of mixture components.
19.	w_i	component	The i^{th} component in a k-vector of constraints placed on the z_i $(1 \leq i \leq k)$ where the z_i are constrained variables. Box and Gardner [5].
20.	x_i, ξ_i	mixture components	Represents the proportion of the i^{th} $(1 \leq i \leq k)$ component or constituent in a mixture.
21.	x_{ij}		The proportion of the j^{th} $(1 \leq j \leq n_i)$ component in the i^{th} $(1 \leq i \leq n)$ "major" component present in a mixture.
22.	y_u	observed response	The response observed by the experimenter in the u^{th} experiment or at the u^{th} setting in the design matrix of the concomitant variables.
23.	$\hat{y}(\cdot)$	predicted response	The value of the predicted response at a particular point in the factor space.
24.	z_i		The i^{th} constrained variable. Box and Gardner [5].

Part II: Chapters 3 to 8.

	Symbol	Term	Definition
1.	$\underset{\sim}{A}$	matrix	The alias matrix, $\underset{\sim}{A}=(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{W}_2$. This matrix is used to show the bias in the estimates of the parameters in a first-degree model if the true polynomial function is of higher degree than 1.
2.	a_i	radius multiplier	The radius multiplier for Simplex-sum designs.
3.	$\underset{\sim}{\alpha}_1$, $\underset{\sim}{\alpha}_2$	unknown parameters	Vectors of unknown parameters used in the models of first and second degree respectively.
4.	$\underset{\sim}{\alpha}_3$		Vector of unknown parameters used for measuring the effects of the independent design variables.
5,	$\underset{\sim}{\alpha}_4$		Vector of unknown parameters used for measuring the contrast effects. When only one contrast is considered, $\underset{\sim}{\alpha}_4=\alpha_{L_1}$.
6.	β_i, γ_i		Part I (1.)
7.	B	bias	The square of the bias of $\hat{y}$ averaged over the region of interest.

	Symbol	Term	Definition
8.	$b_i, c_i, \hat{\alpha}_i, \hat{\beta}_i, \hat{\gamma}_i$		Part I (2.)
9.	$\underset{\sim}{C}_1$	matrix	An N x (q-1) matrix the elements of whose columns are the coefficients of the q-1 contrasts of the observations.
10.	$\underset{\sim}{C}$		An N x q matrix: $\underset{\sim}{C} = [\underset{\sim}{j}_N \; \underset{\sim}{C}_1]$.
11.	c	radius multiplier	Radius multiplier which when multiplied by the elements in a row of the design matrix defines the radius of the design as $c\sqrt{p}$, where p is the number of design variables.
12.	$\underset{\sim}{D}_w$, $\tilde{\underset{\sim}{W}}$	design matrix	An N x (k-q) matrix of design variables whose elements represent the levels of the design variables to be run in the N experiments. It has rank k-q.
13.	$\underset{\sim}{D}_x$		An N x k design matrix of the mixture components.
14.	$\underset{\sim}{D}$		A (q-1) x (q-1) matrix used in the definition of the moment matrix for comparing the method of orthogonal contrasts to the method of orthogonal blocking.

	Symbol	Term	Definition
15.	$\underset{\sim}{D}^{(i)}$		The i^{th} subset of design points used in the design matrix of the Simplex-sum designs.
16.	$\underset{\sim}{D}_{wl}$		Upper or lower half of the design matrix.
17.	$\underset{\sim}{D}_1, \underset{\sim}{D}_2, \underset{\sim}{D}_3$		Matrices, which together, comprise the design matrix of a central composite design.
18.	$\underset{\sim}{\Delta}$		A matrix used for defining the quantities which make-up the average square bias formula.
19.	$\eta(\cdot)$	response	Part I (7.)
20.	ϵ_u	error	Part I (6.)
21.	$\underset{\sim}{F}$	matrix	A matrix which is used in defining the design matrix of the independent design variables.
22.	$f(\cdot)$	function	Part I (8.)
23.	G	moment	The pure fourth moment of a second-degree design.
24.	h_i	scalar quantity	A scalar quantity used for defining the symmetric interval of interest for the i^{th} $(1 \leq i \leq k)$ mixture component.

	Symbol	Term	Definition
25.	$\underset{\sim}{H}$	matrix	A diagonal matrix of scale constants. The j^{th} $(1 \leq j \leq k)$ element in the diagonal is h_j. This matrix is k x k.
26.	$\underset{\sim}{I}_p$		An identity matrix of order p.
27.	J	scalar quantity	Represents the mean square error of $\hat{y}$ averaged over the region of interest.
28.	$\underset{\sim}{1}_\ell$	vector	An ℓ x 1 column vector of ones.
29.	k	scalar quantity	Part I (10.)
30.	K		Scalar quantity used in the formula for the average variance of $\hat{y}$.
31.	L_i	contrast	The i^{th} contrast of the observations.
32.	M		The number of rows in the factorial portion of the design matrix for a second-degree central composite design.
33.	μ		Part I (15.)
34.	$\underset{\sim}{\mu}_{\ell f}$	matrix	Matrix of region moments derived by integrating the product of the terms in the models of degree ℓ, f over the region of interest.

	Symbol	Term	Definition
35.	n_j		Part I (13.)
36.	N	number	Part I (16.)
37.	N_ω		Number of design points in the axial group. $N_\omega=2p+n_{0\omega}$ where $n_{0\omega}$ is the number of center point replicates run while performing the axial group.
38.	N		Part I (17.)
39.	N_j		The number of rows in the j^{th} $(1\leq j\leq q)$ compartment of the design matrix. The number of observations in the j^{th} $(1\leq j\leq q)$ group.
40.	λ_i, [ii]	moment	Moments of a design.
41.	Ω		The area (volume) of the region of interest. This quantity is used in the formula for J.
42.	P_i		Quantity used in defining the moment matrix of a second-degree design.
43.	q		Part I (18.)
44.	ρ^*	radius	The radius of the largest sphere centered at the point of interest that will fit inside the convex polytope.

	Symbol	Term	Definition
45.	R*	region	Extended region of interest. This region has radius ρ^*.
46.	R		Region of interest. This region is ellipsoidal in the system of the mixture components and is a unit sphere in the system of the design variables.
47.	S_i		The sum of the number of mixture components up through the i^{th} category, e.g., $S_2=n_1+n_2$.
48.	$\underset{\sim}{T}$	transformation matrix	A k x k orthogonal matrix used in the transformation from the k dependent intermediate variables to k-q independent design variables together with q zeroes.
49.	$\underset{\sim}{T}_1$		A k x (k-q) partition of the matrix $\underset{\sim}{T}$. The matrix $\underset{\sim}{T}_1$ contains k-q orthonormal columns.
50.	$\underset{\sim}{T}_2$		A k x q partition of the matrix $\underset{\sim}{T}$, the elements of whose columns are such that $\underset{\sim}{V}\ \underset{\sim}{T}_2 = \underset{\sim}{0}$.
51.	t_{ij}	element	The element in the i^{th} row and j^{th} column of the matrix $\underset{\sim}{T}$.
52.	v_i	variable	The i^{th} $(1\leq i\leq k)$ intermediate variable.

	Symbol	Term	Definition
53.	$\underset{\sim}{V}$	matrix	An N x k matrix the elements of which are the levels of the intermediate variables to be used in the N experiments. It has rank k-q.
54.	V	scalar quantity	The variance of $\hat{y}$ averaged over the region of interest.
55.	w_i	variable	The i^{th} $(1 \leq i \leq k-q)$ design variable.
56.	$\underset{\sim}{W}$	matrix	An N x k matrix formed by adding q columns, the elements of which are not all zero, onto the matrix $\underset{\sim}{\tilde{W}}$. The matrix $\underset{\sim}{W}$ has rank k.
57.	$\underset{\sim}{W}_1$		An N x k matrix whose rows contain the levels of the variables in a first-degree model for the N experiments.
58.	$\underset{\sim}{W}_2$		An N x $\binom{p}{2}$ matrix whose columns represent the terms in a second-degree polynomial not contained in a polynomial of degree 1.
59.	x_i	component	Part I (20.)
60.	$\underset{\sim}{x}_0$	base point	A vector of k components each of which specifies the center

	Symbol	Term	Definition
			of the interval of interest for the corresponding mixture component. This vector specifies the location of the base point in the factor space of the mixture components.
61.	$\underset{\sim}{X}$	matrix	An N x k matrix containing the elements $x_{ui}-x_{0i}$ $(1\leq u\leq N,\ 1\leq i\leq k)$. The matrix $\underset{\sim}{X}$ has rank k-q.
62.	y_u	response	Part I (22.)
63.	$y(.)$		Part I (23.)
64.	Y_j		The sum of the n observations in the j^{th} $(1\leq j\leq q)$ group.
65.	$\bar{y}_j$		The mean of the observations in the j^{th} $(1\leq j\leq q)$ group.
66.	z_{iu}	contrast coefficient	The u^{th} $(1\leq u\leq N)$ coefficient in the i^{th} contrast of the observations.
67.	$z_{i.}^{(j)}$		The values of the coefficients in the j^{th} group $(1\leq j\leq q)$ in the i^{th} contrast. All $z_{iu}^{(j)}$ are equal in a particular group.
68.	ω		Level of the axial points in a central composite design.

APPENDIX B.

THE GEOMETRY OF n DIMENSIONS: SECTIONS AND FRUSTA OF SIMPLICES.

The elements of the series {point, line, plane, hyperplane, ..., n-flat} are regions determined by 1, 2,..., n+1 points and have 0, 1, 2, ..., n dimensions. In n dimensions, if we have p ($p \leq n$) equations which are linear in the variables w_i ($1 \leq i \leq n$), then the p equations define an (n-p)-flat. That is, a p-flat is a flat space of p dimensions and is determined by any p+1 points. Also, every r-flat ($r \leq p$) which is determined by r+1 of the p+1 points lies entirely in the p-flat.

Let the p+1 points uniquely define a p-flat. These p+1 points then cannot be contained in the same (p-1)-flat. Also, no q of them ($q \leq p$) can be contained in the same (q-2)-flat, for if they do, then this (q-2)-flat, which is determined by q-1 points, together with the remaining (p+1)-q points would determine a (p-1)-flat since (p+1)-q+q-1 = p points. Thus a system of any p+1 points, no q of which lie in the same (q-2)-flat is a system of linearly independent points.

A polytope, the analogue of a polygon in two dimensions and a polyhedron in three dimensions, is a figure bounded by hyperplanes (flats). In n dimensions, the figure is bounded by a set of (n-1)-flats. Adjacent (n-1)-flats meet in boundaries of (n-2)-dimensions or (n-2)-flats and in

the types of polytopes that we are considering, two and only two (n-1)-flats will meet in a boundary of (n-2)-dimensions. A polytope is <u>convex</u> when it lies entirely to one side of each of its (n-1)-boundaries.

The least number of (n-1)-flats which can enclose a space and form a polytope is n+1. Therefore in two dimensions, we have the triangle. In three dimensions, the tetrahedron. Such a figure is called a simplex. A simplex of n dimensions has $\binom{n+1}{r+1}$ boundaries of r dimensions.

Let us now consider the two category mixture situation. The reason for considering two categories is primarily because figures in two and three dimensions can be visualized, and with two categories when k=4, 5, these figures can be drawn and intuitively apprehended. When n is the number of components in category 1 and k ($k \geq n+2$) is the total number of components, the <u>section</u> of the simplex is that part of the simplex specified by the constraint

$$\sum_{i=1}^{n} x_i = \sum_{j=n+1}^{k} x_j = \frac{1}{2} . \qquad \text{(B1)}$$

Since we are considering the interior as well as the boundaries of the simplex, we see from (B1) that the section will be a polytope of dimensionality 1 less that that of the simplex.

The other polytopes formed by cutting the simplex with the section (B1) are called <u>frusta</u> of the simplex. The

forms of the section and the frusta depend on the position of the hyperplane (B1) with regard to the vertices of the simplex and Sommerville [21] points out that there is a close connection between the form of the section and those of the frusta.

From (B1), we see that the hyperplane separates the k vertices of the simplex into two groups, one with n vertices and the other with k-n vertices. Sommerville [21] denotes the type of polytope of the section by (n, k-n) so that of course, (n, k-n) = (k-n, n). The other types of polytopes, the frusta are denoted by (n|k-n) and (k-n|n), where the first of the two numbers in the symbol represents the number of vertices of the simplex which the frustum contains.

Let us first consider a tetrahedron and denote the vertices by A, B, C, D as in Figures 4a and 4b on the next page. Let the section divide the vertices into groups 2, 2. The section (2, 2) is a quadrilateral (Plane E in Figure 4a), and the two frusta are of the same type (2|2), pentahedra which are bounded by two triangles and three quadrilaterals.

Now consider the case where the plane of section divides the vertices into groups 1, 3. This does not conform to the constraint (B1) since n must be greater than 2 but does help clarify the subject of sections and frusta. Then the section is a triangle (PQR in Figure 4b), i.e., (1, 3) denotes a triangle and the frusta are of the type (1|3), a tetrahedron and (3|1), a pentahedron which is bounded by two triangles and three quadrilaterals. This

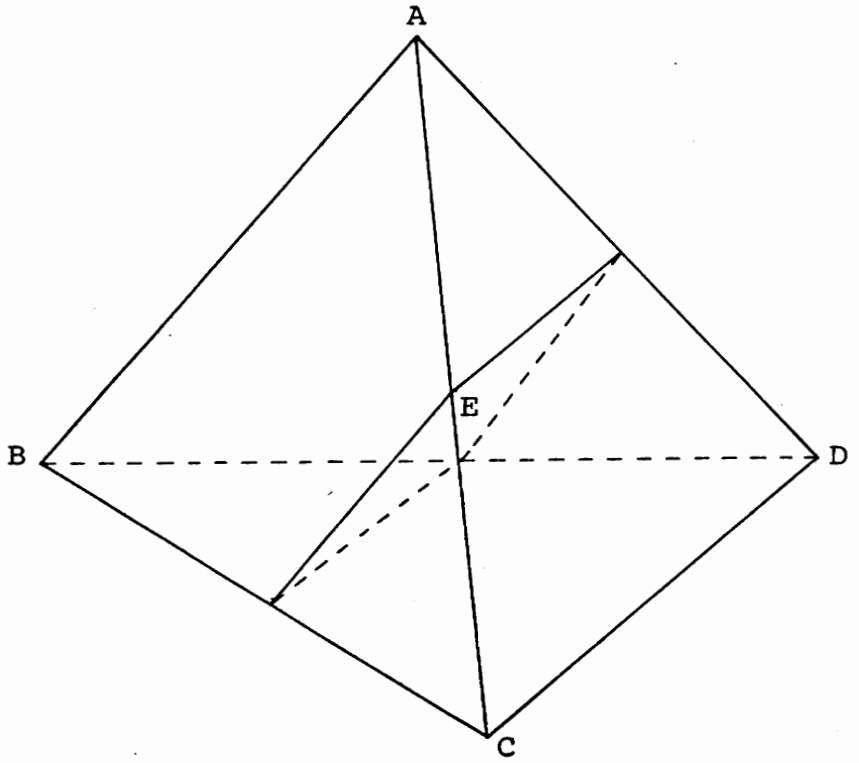

Figure 4a.

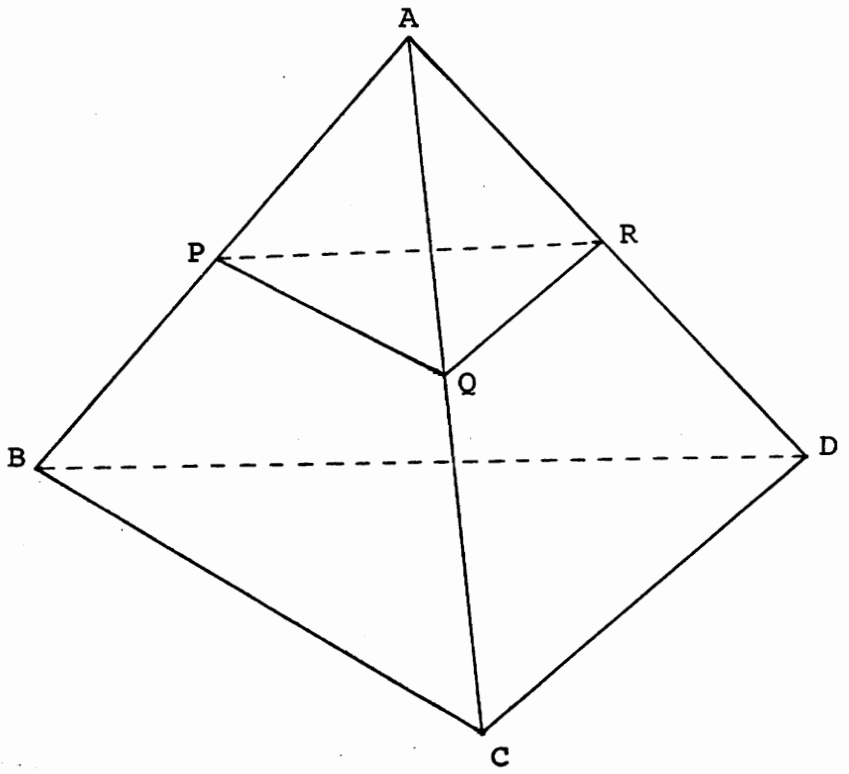

Figure 4b.

Sections and Frusta of the Simplex ABCD.

latter example corresponds to the two category situation outlined in Section 1.3 where category 1 has one component and category 2 has k'=3 components.

Finally let us consider the simplex ABCDE and let the hyperplane divide the vertices into groups 2, 3 say AB and CDE. According to Sommerville [21], the vertices of the section are the points of intersection of the lines AC, AD, AE, BC, BE. The edges are the lines of intersection of the planes ABC, ABD, ABE, ACD, ACE, ADE, BCD, BCE, BDE; and the faces are the planes of intersection of the hyperplanes ABCD, ABCE, ABDE, ACDE, BCDE. The first three of these faces are (2, 2) sections, i.e., quadrilaterals, and the last two are (1, 3) sections, i.e., triangles. Thus the section is a pentahedron (3|1) or (2|2) and we have the description of the factor space corresponding to the case involving two categories where n=2 or n=3 and k=5.

APPENDIX C.

TO SHOW THAT THE DIMENSIONALITY OF THE BOUNDARY CLOSEST TO ANY INTERIOR POINT OF A CONVEX POLYTOPE IS 1 LESS THAN THE DIMENSIONALITY OF THE POLYTOPE.

In Appendix B, we defined a body or polytope of n dimensions as being convex if it lies entirely to one side of each of its (n-1)-boundaries. That is, if P_1 and P_2 are any two points in or on the body, so also are all the points of the interval P_1P_2. Generalizing then, all interior angles which are formed by the intersection of any two adjacent boundaries of the polytope must be less than two right angles.

Suppose the shortest distance from an interior point O is to a point A on the boundary hyperplane which we denote by Π. Then the interior angle between OA and Π is $\geq 90^{\circ}$. For suppose it is not, i.e., let the interior angle between OA and Π be less than 90°. Then there exists a point A' close to A such that OA'<OA.

Now suppose that the shortest distance from the point O to a boundary is described as the length of the projection OB where the point B is on the intersection of two boundary hyperplanes Π_1 and Π_2. Let the hyperplanes Π_1 and Π_2 each be of dimensionality n-1. Then the intersection is of dimensionality n-2. Now the interior angle between OB and Π_1 is $\geq 90^{\circ}$, i.e., angle $OBC_1 \geq 90^{\circ}$, where C_1 is any point in Π_1. Also, the interior angle between OB and Π_2 is $\geq 90^{\circ}$, i.e.,

angle $OBC_2 \geq 90^\circ$ where C_2 is <u>any</u> point in Π_2. Therefore, if we take any plane through OB, it will interesct Π_1 and Π_2 in two lines BC_1 and BC_2 and such that angle $OBC_1 \geq 90^\circ$ and angle $OBC_2 \geq 90^\circ$. Hence the angle $C_1BC_2 \geq 180^\circ$ and we have a contradiction to the statement that all interior angles of a convex polytope are less than two right angles. Therefore, the shortest distance from an interior point is not to a boundary of dimensionality n-2.

APPENDIX D.

THE MOMENTS OF A ROTATABLE DESIGN.

The class of rotatable designs is characterized by the property that these designs generate constant information (in terms of the variance of the predictor $\hat{y}$) on each sphere whose center is the center of the designs. In other words, the variance (inverse of the information) of the predictor $\hat{y}$ at some point in the design space is a function only of the distance the point lies from the center of the design and is not a function of the direction.

The concept of rotatability for designs of degree greater than 1 was first introduced by Box and Hunter [6]. Box and Hunter show that the necessary and sufficient conditions for a design to be d^{th} degree rotatable is that the generating function of the moments up to degree 2d, given by

$$Q = \frac{1}{N}\sum_{u=1}^{N}(1 + t_1 w_{u1} + t_2 w_{u2} + \ldots + t_p w_{up})^{2d}, \quad (D1)$$

should be of the form

$$Q = \sum_{s=0}^{d} a_{2s}(t_1^2 + t_2^2 + \ldots + t_p^2)^s \quad (D2)$$

where a_{2s} (s=0, 1, ..., d) are constants depending on the design points and are independent of the t_i, and the w_{ui} ($1 \leq i \leq p$) are the levels of the independent design variables. That is to say, if the moment

$$\frac{1}{N}\sum_{u=1}^{N} w_{u1}^{\theta_1} w_{u2}^{\theta_2} \cdots w_{up}^{\theta_p}$$

is denoted by

$$[1^{\theta_1}\, 2^{\theta_2} \cdots p^{\theta_p}],$$

then by equating the coefficients of $t_1^{\theta_1} t_2^{\theta_2} \cdots t_p^{\theta_p}$ in (D1) and (D2), Box and Hunter show that for a set of points to be d^{th} degree rotatable, it is necessary and sufficient that the design moments satisfy the condition that

$$[1^{\theta_1}\, 2^{\theta_2} \cdots p^{\theta_p}] = 0 \qquad \text{if all or any } \theta_i \text{ are odd}$$

$$= \frac{\lambda_\theta \prod_{i=1}^{p} \theta_i!}{2^{\frac{\theta}{2}} \prod_{i=1}^{p} \left(\frac{\theta_i}{2}\right)!} \qquad \text{if all } \theta_i \text{ are even.} \tag{D3}$$

In (D3), $\theta = \theta_1 + \theta_2 + \ldots + \theta_p \leq 2d$ and λ_θ is a constant depending on θ but is unaffected by the way in which θ is partitioned into $\theta_1, \theta_2, \ldots, \theta_p$.

In this paper, since we are concerned with designs of degree 1 and 2, we shall show, using (D3), the method for constructing the moment matrix associated with first and second-degree rotatable designs. For d = 1, the only non-zero moment for a rotatable design is $[i^2]$, and therefore from (D3),

$$[i^2] = \frac{\lambda_2 \cdot 2!}{2 \cdot 1}$$

$$= \lambda_2 . \tag{D4}$$

Hence the moment matrix for a first-degree rotatable design where $\underset{\sim}{w}_u' = (1, w_{u1}, w_{u2}, \ldots, w_{up})$ is

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' \\ \underset{\sim}{0} & \lambda_2 \underset{\sim}{I}_p \end{bmatrix} . \qquad \text{(D5)}$$

For d=2, the non-zero moments are $[i^2]$, $[i^2j^2]$, and $[i^4]$. Hence from (D3),

$$[i^2j^2] = \frac{\lambda_4\, 2!\, 2!}{2^2\, 1!1!}$$

$$= \lambda_4 \qquad i \neq j, \qquad \text{(D6)}$$

and

$$[i^4] = \frac{\lambda_4\, 4!}{2^2 \cdot 2!}$$

$$= 3\lambda_4 . \qquad \text{(D7)}$$

Using (D4), (D6) and (D7) then, we have for the moment matrix of a second-degree rotatable design where

$$\underset{\sim}{w}_u' = (1, w_{u1}, \ldots, w_{up}, w_{u1}^2, \ldots, w_{up}^2, w_{u1}w_{u2}, \ldots, w_{u,p-1}w_{up}),$$

$$\frac{1}{N}\underset{\sim}{W}'\underset{\sim}{W} = \begin{bmatrix} 1 & \underset{\sim}{0}' & \lambda_2 \underset{\sim}{j}_p' & \underset{\sim}{0}' \\ \underset{\sim}{0} & \lambda_2 \underset{\sim}{I}_p & \underset{\sim}{0} & \underset{\sim}{0} \\ \lambda_2 \underset{\sim}{j}_p & \underset{\sim}{0} & \lambda_4(2\underset{\sim}{I}_p + \underset{\sim}{j}_p\underset{\sim}{j}_p') & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \lambda_4 \underset{\sim}{I}_{\binom{p}{2}} \end{bmatrix} \qquad \text{(D8)}$$

where $\underset{\sim}{j}_p$ is a p x 1 column vector of ones.

APPENDIX E.

THE CONSTRUCTION OF AN ORTHOGONAL MATRIX $\underset{\sim}{T}$ FOR GENERAL q.

In Chapter 3, the k x k orthogonal matrix $\underset{\sim}{T}$ was partitioned into the two matrices $\underset{\sim}{T}_1$ and $\underset{\sim}{T}_2$. The matrix $\underset{\sim}{T}_1$ was a k x (k-q) matrix containing k-q orthonormal columns and $\underset{\sim}{T}_2$ was a k x q matrix such that $\underset{\sim}{V}\underset{\sim}{T}_2=\underset{\sim}{0}$, where $\underset{\sim}{V}$ is an N x k matrix containing the levels of the intermediate variables. The elements of the matrices $\underset{\sim}{T}_1$ and $\underset{\sim}{T}_2$ were given for q=2.

In this appendix, we show that given a form for general q of the matrix $\underset{\sim}{T}_2$ so that $\underset{\sim}{V}\underset{\sim}{T}_2=\underset{\sim}{0}$, a corresponding form of the k x (k-q) matrix $\underset{\sim}{T}_1$ can be derived so that the k x k matrix $\underset{\sim}{T}$ is orthogonal.

Let us define the quantity

$$S_i = \sum_{\ell=1}^{i} n_\ell \qquad \text{(E1)}$$

where n_ℓ $(1\leq\ell\leq q)$ is the number of mixture components belonging to the ℓ^{th} category. Let

$$h^{j*} = \left[\sum_{\ell=S_{j-1}+1}^{S_j} h_\ell^2\right]^{\frac{1}{2}} \qquad (1\leq j\leq q) \qquad \text{(E2)}$$

and

$$\tilde{h}_i = \frac{h_i}{h^{j*}} \qquad (S_{j-1}+1\leq i\leq S_j) \qquad \text{(E3)}$$

be defined where h_i $(S_{j-1}+1\leq i\leq S_j)$ is the quantity used in the transformation (3.1.2), that is,

$$v_i = \frac{x_i - x_{0i}}{h_i} \qquad (S_{j-1}+1 \le i \le S_j ; 1 \le j \le q).$$

If we define the matrix $\underset{\sim}{T}$ to be of the form

$$\underset{\sim}{T} = [\, \tilde{\underset{\sim}{h}}, \underset{\sim}{T}_1, \underset{\sim}{h}^{(1)}, \underset{\sim}{h}^{(2)}, \ldots, \underset{\sim}{h}^{(q-1)} \,] \qquad \text{(E4)}$$

where $\underset{\sim}{T}_1$ is a k x (k-q) matrix containing k-q orthonormal columns, $\tilde{\underset{\sim}{h}}$ is a k x 1 column vector with the elements

$$\tilde{\underset{\sim}{h}}' = (\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_{S_1}, 0, 0, \ldots, 0),$$

and $\underset{\sim}{h}^{(i)}$ $(1 \le i \le q-1)$ is a column vector containing $S_{i+1}-S_i$ non-zero elements, then one particular construction of the matrix $\underset{\sim}{T}$ is

$$\underset{\sim}{T} = \begin{bmatrix} \tilde{\underset{\sim}{h}} & [\,] & \underset{\sim}{0} & \underset{\sim}{0} & \ldots & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{0} & \ldots & \underset{\sim}{0} \\ \underset{\sim}{0} & [\,] & \underset{\sim}{0} & \ldots & \underset{\sim}{0} & \underset{\sim}{h}^{(1)} & \underset{\sim}{0} & \ldots & \underset{\sim}{0} \\ \underset{\sim}{0} & \underset{\sim}{0} & [\,] & & \underset{\sim}{0} & \underset{\sim}{0} & \underset{\sim}{h}^{(2)} & & \underset{\sim}{0} \\ \vdots & \vdots & & & \vdots & \vdots & & & \vdots \\ \underset{\sim}{0} & \underset{\sim}{0} & \ldots & \underset{\sim}{0} & [\,] & \underset{\sim}{0} & \underset{\sim}{0} & \ldots\ \underset{\sim}{0} & \underset{\sim}{h}^{(q-1)} \end{bmatrix}. \qquad \text{(E5)}$$

The non-zero elements in the last q-1 columns of the matrix $\underset{\sim}{T}$ are the elements in the column vectors $\underset{\sim}{h}^{(i)}$ $(1 \le i \le q-1)$, which we shall define shortly. With the form (E5) of the matrix $\underset{\sim}{T}$, then

$$\underset{\sim}{V}\underset{\sim}{T} = [\underset{\sim}{0}\ \tilde{\underset{\sim}{W}}\ \underset{\sim}{0}\ \underset{\sim}{0}\ \ldots\ \underset{\sim}{0}]$$

where the form of the matrix $\underset{\sim}{V}$ is described at the beginning of this appendix.

From (E5), we see that the matrix $\underset{\sim}{T}_1$ is a diagonal matrix of matrices. In fact, we shall show that each matrix

in $\underset{\sim}{T}_1$ is an n_ℓ x $(n_\ell-1)$ matrix where n_ℓ is defined in (E1).

Denote by t_{ij} the element in the i^{th} row and j^{th} column of the matrix $\underset{\sim}{T}$. From (E4), we see that the first column of the matrix $\underset{\sim}{T}$, that is, t_{i1} $(1\leq i\leq k)$ is

$$\begin{aligned} t_{i1} &= \tilde{h}_i \\ &= \frac{h_i}{h^{1*}} \quad \text{for } i=1,2,\ldots,S_1 \\ &= 0 \quad \text{for } S_1+1\leq i\leq k \end{aligned} \tag{E6}$$

where S_i and h^{i*} $(1\leq i\leq q)$ are defined in (E1) and (E2) respectively. Now, in the ℓ^{th} matrix in the diagonal of the matrix $\underset{\sim}{T}_1$, the elements are

$$t_{ij} = h_{i+1} \quad \begin{matrix} i=S_{\ell-1}+1 \\ j=S_{\ell-1}+3-\ell \end{matrix} \qquad t_{ij} = h_i h_{j+\ell-1} \quad \begin{matrix} i=S_{\ell-1}+1,\ldots,j+\ell-2 \\ j=S_{\ell-1}+4-\ell,\ldots,S_\ell-\ell+1 \end{matrix}$$

$$t_{ij} = -h_{i-1} \quad \begin{matrix} i=S_{\ell-1}+2 \\ j=S_{\ell-1}+3-\ell \end{matrix} \qquad t_{ij} = 0 \quad \begin{matrix} i=j+\ell,\ldots,S_\ell \\ j=S_{\ell-1}+3-\ell,\ldots,S_\ell-\ell \end{matrix}$$

$$t_{ij} = -\Big(\sum_{e=S_{\ell-1}+1}^{j+\ell-2} h_e^2\Big) \quad \begin{matrix} i=j+\ell-1 \\ j=S_{\ell-1}+4-\ell,\ldots,S_\ell-\ell+1. \end{matrix} \tag{E7}$$

Column-normalizing the elements in (E7) gives the matrix $\underset{\sim}{T}_1$.

Let us look now at the last q-1 columns of the matrix $\underset{\sim}{T}$. The columns consist of the n_{j+1} x 1 vectors $\underset{\sim}{h}^{(j)}$ $(1\leq j\leq q-1)$, where the vector $\underset{\sim}{h}^{(j)}$ is in the $k-q+j+1^{st}$ column of the matrix $\underset{\sim}{T}$ (from the left). In order to make these q-1 columns orthogonal to the columns of the matrix $\underset{\sim}{T}_1$ and at the same time insure that the dot product of the elements of $\underset{\sim}{h}^{(j)}$ equals 1, that is, $\underset{\sim}{h}^{(j)\prime}\underset{\sim}{h}^{(j)}=1$, the elements of the vector

$\underset{\sim}{h}^{(j)}$ will be denoted by $\tilde{h}_f$ where

$$\tilde{h}_f = \frac{h_f}{h^{j+1*}} \qquad (S_j+1 \le f \le S_{j+1}) \tag{E8}$$

and

$$h^{j+1*} = \left[\sum_{e=S_j+1}^{S_{j+1}} h_e^2 \right]^{\frac{1}{2}} . \tag{E9}$$

For example, the second matrix ($\ell=2$) down the diagonal of the matrix $\underset{\sim}{T}_1$ (before column-normalizing) is of the form

$$\begin{bmatrix}
h_{S_1+2} & h_{S_1+1}h_{S_1+3} & h_{S_1+1}h_{S_1+4} & \cdots & h_{S_1+1}h_{S_2} \\
-h_{S_1+1} & h_{S_1+2}h_{S_1+3} & h_{S_1+2}h_{S_1+4} & \cdots & h_{S_1+2}h_{S_2} \\
0 & -\sum_{j=1}^{2} h_{S_1+j}^2 & h_{S_1+3}h_{S_1+4} & \cdots & h_{S_1+3}h_{S_2} \\
0 & 0 & -\sum_{j=1}^{3} h_{S_1+j}^2 & & h_{S_1+4}h_{S_2} \\
0 & 0 & 0 & & \\
\vdots & \vdots & \vdots & & \vdots \\
 & & & & \\
0 & 0 & 0 & & -\sum_{j=1}^{S_2-S_1-1} h_{S_1+j}^2
\end{bmatrix} , \tag{E10}$$

and the corresponding elements in the vector $\underset{\sim}{h}^{(1)}$ are

$$\underset{\sim}{h}^{(1)} = \begin{bmatrix} \tilde{h}_{S_1+1} \\ \tilde{h}_{S_1+2} \\ \vdots \\ \tilde{h}_{S_2} \end{bmatrix} .$$

APPENDIX F.

THE GRAM-SCHMIDT ORTHOGONALIZATION PROCESS.

Let the matrix $\underset{\sim}{T}$ be of the form (E5) in Appendix E. With the ℓ^{th} matrix in the diagonal of the matrix $\underset{\sim}{T}_1$ and the associated vector $\underset{\sim}{h}^{(\ell-1)}$, we want to construct an orthogonal basis of dimensionality n_ℓ which is a subspace of the k-dimensional vector space spanned by the columns of the matrix $\underset{\sim}{T}$.

Let $\ell=1$ and choose the vector $\underset{\sim}{X}_1 = \underset{\sim}{\tilde{h}}$ where

$$\underset{\sim}{\tilde{h}}' = (\tilde{h}_1, \tilde{h}_2, \ldots, \tilde{h}_{n_1}, 0, 0, \ldots, 0).$$

The Gram-Schmidt orthogonalization process consists of constructing the vectors $\underset{\sim}{Y}_i$ $(1 \leq i \leq n_1)$ in the following manner,

$$\begin{aligned}
\underset{\sim}{Y}_1 &= \underset{\sim}{X}_1 \\
\underset{\sim}{Y}_2 &= \underset{\sim}{X}_2 - \frac{\underset{\sim}{Y}_1' \underset{\sim}{X}_2}{\underset{\sim}{Y}_1' \underset{\sim}{Y}_1} \underset{\sim}{Y}_1 \\
&\vdots \\
\underset{\sim}{Y}_{n_1} &= \underset{\sim}{X}_{n_1} - \frac{\underset{\sim}{Y}_{n_1-1}' \underset{\sim}{X}_{n_1}}{\underset{\sim}{Y}_{n_1-1}' \underset{\sim}{Y}_{n_1-1}} \underset{\sim}{Y}_{n_1-1} - \cdots - \frac{\underset{\sim}{Y}_{n_1}' \underset{\sim}{X}_{n_1}}{\underset{\sim}{Y}_1' \underset{\sim}{Y}_1}
\end{aligned} \qquad \text{(F1)}$$

so that the vectors

$$\underset{\sim}{G}_i = \frac{\underset{\sim}{Y}_i}{\| \underset{\sim}{Y}_i \|_2^1} \qquad i=1,2,\ldots,n_1 \qquad \text{(F2)}$$

form an orthogonal (orthonormal) basis of dimensionality n_1. For example, let $n_1=3$ and $\underset{\sim}{X}_1' = (h_1, h_2, h_3)$. Put $\underset{\sim}{X}_2' = (1,0,0)$ and $\underset{\sim}{X}_3' = (0,1,0)$. Then from (F1), we have

$$\underset{\sim}{Y}_1 = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

$$\underset{\sim}{Y}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} - \frac{[h_1 \; h_2 \; h_3]\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}}{[h_1 \; h_2 \; h_3]\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

$$= \begin{bmatrix} h_2^2 + h_3^2 \\ -h_1 h_2 \\ -h_1 h_3 \end{bmatrix} \frac{1}{h_1^2+h_2^2+h_3^2}$$

$$\underset{\sim}{Y}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - \frac{\underset{\sim}{Y}_2' \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{\underset{\sim}{Y}_2' \underset{\sim}{Y}_2} \underset{\sim}{Y}_2 - \frac{[h_1 \; h_2 \; h_3]\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}}{h_1^2+h_2^2+h_3^2} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$$

$$= \begin{bmatrix} 0 \\ h_3^2 \\ -h_2 h_3 \end{bmatrix} \frac{1}{h_2^2+h_3^2} \; .$$

Then the upper left-hand 3 x 2 matrix in the k x k matrix $\underset{\sim}{T}$ is

$$\begin{bmatrix} \frac{\underset{\sim}{Y}_2}{D_1} & \frac{\underset{\sim}{Y}_3}{D_2} \end{bmatrix}$$

where

$$D_1 = \left[\frac{(h_2^2 + h_3^2)}{h_1^2 + h_2^2 + h_3^2} \right]^{\frac{1}{2}}$$

$$D_2 = \left[\frac{h_3^2}{h_2^2 + h_3^2} \right]^{\frac{1}{2}} \; .$$

APPENDIX G.

A DISCUSSION OF THE PROCEDURE FOR MINIMIZING THE AVERAGE MEAN SQUARE ERROR OF $\hat{y}$ WHEN NEITHER THE VARIANCE NOR THE BIAS OF $\hat{y}$ CAN BE IGNORED.

In this appendix, we attempt to reach some compromise in the practical situation when neither the contribution from the variance of $\hat{y}$ nor the contribution of the bias of $\hat{y}$ can be ignored. In other words, we shall attempt to minimize the quantity $J = V + B$ where V is the average variance of $\hat{y}$ and B is the average squared bias of $\hat{y}$, where average means averaged over the region of interest. The region of interest will be the extended region of interest R*.

It has been shown ([3], Appendix 1) that the average squared bias B can be written as

$$B = \frac{N}{\sigma^2} \underset{\sim}{\alpha}_2' \underset{\sim}{\Delta} \underset{\sim}{\alpha}_2 \qquad (G1)$$

where $\underset{\sim}{\Delta}=\underset{\sim}{A}'\underset{\sim}{\mu}_{11}\underset{\sim}{A}-\underset{\sim}{\mu}_{12}'\underset{\sim}{A}-\underset{\sim}{A}'\underset{\sim}{\mu}_{12}+\underset{\sim}{\mu}_{22}$, $\underset{\sim}{\alpha}_2$ is the vector of parameters associated with the second-degree terms in a second-degree model, $\underset{\sim}{A}$ is the alias matrix $(\underset{\sim}{W}_1'\underset{\sim}{W}_1)^{-1}\underset{\sim}{W}_1'\underset{\sim}{W}_2$ and $\underset{\sim}{\mu}_{ef}$ (e,f=1,2) are the region moment matrices. When q=2, the expression (5.3.2) in Chapter 5 for the average variance of $\hat{y}$ was shown to be

$$V = 2 + \left(\frac{p}{p+2}\right)\frac{\rho^{*2}}{\lambda_2}$$

where p is the number of independent variables (p=k-2) and ρ^* is the radius of the region of interest. Now, if we

combine this last expression for V with the expression (G1) for the average squared bias to form J=V+B, then it can be shown that for any first-degree rotatable design in which the third-degree moments vanish,

$$J = \left\{2 + \left(\frac{p}{p+2}\right)\frac{\rho^{*2}}{\lambda_2}\right\} + \left\{\lambda_2 - \frac{\rho^{*2}}{p+2}\right\}^2 \left(\sum_{i=1}^{p} \tilde{\alpha}_{ii}\right)^2$$

$$+ \frac{2\rho^{*4}(p+2)\sum_{i=1}^{p} \tilde{\alpha}_{ii}^2 + \rho^{*4}(p+2)\sum\sum_{i<j} \tilde{\alpha}_{ij}^2 - 2\rho^{*4}\left(\sum_{i=1}^{p} \tilde{\alpha}_{ii}\right)^2}{(p+2)^2(p+4)} \quad \text{(G2)}$$

where $\tilde{\alpha}_{ij} = \frac{\alpha_{ij}\sqrt{N}}{\sigma}$ is the element in the i^{th} row and j^{th} column of a matrix $[\tilde{\alpha}]$ containing the standardized quadratic parameters. That is, the matrix $[\tilde{\alpha}]$, which is symmetric p x p, is defined [3] as such; the diagonal elements $\tilde{\alpha}_{ii}$ are the parameters associated with the pure quadratic terms multiplied by the quantity $\frac{\sqrt{N}}{\sigma}$, and the off-diagonal elements $\frac{\tilde{\alpha}_{ij}}{2}$ are the crossproduct or interaction parameters of the true surface also multiplied by the quantity $\frac{\sqrt{N}}{\sigma}$.

Let us rewrite J in a simpler form by letting n=p,

$$\theta = \text{trace}\,[\tilde{\alpha}^2] = \sum_{i=1}^{n} \tilde{\alpha}_{ii}^2 + \frac{1}{4}\sum\sum_{i\neq j}^{n} \tilde{\alpha}_{ij}^2$$

$$= \frac{N}{\sigma^2}\left\{\sum_{i=1}^{n} \alpha_{ii}^2 + \frac{1}{2}\sum\sum_{i<j}^{n} \alpha_{ij}^2\right\},$$

and $\Lambda = \frac{(\text{trace}\,[\tilde{\alpha}])^2}{\text{trace}\,[\tilde{\alpha}^2]}$.

Then

$$J = \{2 + \frac{n\rho^{*2}}{(n+2)\lambda_2}\} + \theta\{\Lambda(\lambda_2 - \frac{\rho^{*2}}{n+2}) + \frac{2\rho^{*4}(n+2-\Lambda)}{(n+2)^2(n+4)}\}. \quad (G3)$$

The quantity θ $(0 \leq \theta \leq \infty)$ is defined [3] as an overall measure of the magnitude of the quadratic tendency of the true surface relative to the sampling error and Λ $(0 \leq \Lambda \leq n)$ is defined to be a measure of the condition of the quadratic surface in terms of the latent roots of the matrix of quadratic parameters. It is shown ([22], Appendix B) that both quantities θ, Λ are independent of the elements of the transformation matrix $\underset{\sim}{T}$, and therefore the form (G3) for J can be used in the mixture problem.

Looking at the expression (G3) for J, we see that the value of λ_2 (the second moment of the design) which minimizes J is a function of ρ^{*2}, θ and Λ. Since ρ^{*2} is fixed for a particular situation and if we know θ and Λ (or if estimates $\hat{\theta}$, $\hat{\Lambda}$ are available), the best spread of the design points can be obtained in the following manner. If we set-up J in (G3) as a function of λ_2, we have

$$Q(\lambda_2) = \frac{n\rho^{*2}}{(n+2)\lambda_2} + \theta\Lambda(\lambda_2 - \frac{\rho^{*2}}{n+2}). \quad (G4)$$

To minimize $Q(\lambda_2)$ in (G4), we can take the derivative with respect to λ_2 and set the result equal to zero,

$$\frac{\partial Q(\lambda_2)}{\partial \lambda_2} = -\frac{n\rho^{*2}}{(n+2)\lambda_2^2} + 2\theta\Lambda(\lambda_2 - \frac{\rho^{*2}}{n+2}) = 0.$$

Then

$$2\theta\Lambda\lambda_2^3 - \frac{2\theta\Lambda\rho^{*2}}{n+2}\lambda_2^2 - \frac{n\rho^{*2}}{n+2} = 0,$$

or

$$\lambda_2^3 - \frac{\rho^{*2}}{n+2}\lambda_2^2 - \frac{n\rho^{*2}}{2\theta\Lambda(n+2)} = 0. \qquad \text{(G5)}$$

Now one method of obtaining the roots of the equation (G5) is to reduce (G5) to the form

$$x^3 + ax + b = 0. \qquad \text{(G6)}$$

The equation (G6) has one real root if $c = \frac{b^2}{4} + \frac{a^3}{27} > 0$. This root is

$$x = \{-\frac{b}{2} + c^{1/2}\}^{1/3} + \{-\frac{b}{2} - c^{1/2}\}^{1/3}.$$

Let $\lambda_2 = x + \frac{\rho^{*2}}{3(n+2)}$ and put this value of λ_2 as x in (G6). Then

$$x^3 - \frac{\rho^{*4}}{3(n+2)^2}x - \frac{2\rho^{*6}}{27(n+2)^3} - \frac{n\rho^{*2}}{2\theta\Lambda(n+2)} = 0,$$

which is the form of the equation (G6) where

$$a = -\frac{\rho^{*4}}{3(n+2)^2}, \quad b = -\rho^{*2}\left[\frac{4\rho^{*4}\theta\Lambda + 27n(n+2)^2}{54\theta\Lambda(n+2)^3}\right].$$

If we substitute these values of a and b in the formula for c, we see that the value of c is greater than zero and therefore the value of λ_2 which minimizes (G3) is

$$\lambda_2 = \frac{\rho^{*2}}{3(n+2)} + \{-\frac{b}{2} + c^{1/2}\}^{1/3} + \{-\frac{b}{2} - c^{1/2}\}^{1/3} \qquad \text{(G7)}$$

where in (G7),

$$b = -\rho^{*2}\left[\frac{4\rho^{*4}\theta\Lambda + 27n(n+2)^2}{54\theta\Lambda(n+2)^3}\right],\quad c = \rho^{*4}\left[\frac{8\rho^{*4}\theta\Lambda n + 27n^2(n+2)^2}{432\theta^2\Lambda^2(n+2)^4}\right]$$

assuming θ and Λ are known.

We can now minimize J in (G3) assuming the parameters α_{ii}, α_{ij} $(i<j)$, or functions θ, Λ of the parameters are known. However, since even estimates of these parameters are not usually available, one way to use (G3) is to assume some design configurations and check these designs for robustness (the ability of J to remain fixed for different values of the unknown parameters) for reasonable ranges of values of the unknown parameters. This is the procedure used in [3] and [22], but we shall discuss the method also since the region of interest is not restricted to a unit sphere.

To find the minimizing value of λ_2 for estimates of θ and Λ, we set-up J as a function of the second moment, i.e.,

$$J = V(\lambda_2) + \theta B(\lambda_2) \qquad \text{(G8)}$$

where $V(\lambda_2)$ and $B(\lambda_2)$ are defined in (G3). To determine a reasonable value for λ_2 for a given value of θ, let

$$g = \frac{V(\lambda_2)}{\theta B(\lambda_2)}\ , \qquad \text{(G9)}$$

where g is a positive constant selected by the experimenter for the particular situation. Note that the quantity g is a measure of the ratio of the variance to the bias when J is

minimized for given values of θ and Λ. From [3], for given values of Λ and g then, minimizing values of λ_2 may be found by minimizing the function

$$Q(\lambda_2) = \{V(\lambda_2)\}^g\{B(\lambda_2)\}. \tag{G10}$$

This gives a minimizing value of λ_2 which at the same time satisfies (G9). If we differentiate $Q(\lambda_2)$ in (G10) and set the result equal to zero, we have

$$g\{V(\lambda_2)\}^{g-1}\{B(\lambda_2)\}\{V'(\lambda_2)\} + \{V(\lambda_2)\}^g\{B'(\lambda_2)\} = 0. \tag{G11}$$

If λ_2^* is a solution of (G11) and

$$g\{V'(\lambda_2)\}^{g-1}B(\lambda_2)B'(\lambda_2) \neq 0,$$

then by dividing (G11) by $g\{V'(\lambda_2)\}^{g-1}B(\lambda_2)B'(\lambda_2)$, we have

$$\frac{V(\lambda_2^*)}{gB(\lambda_2^*)} = -\frac{V'(\lambda_2^*)}{B'(\lambda_2^*)}$$

$$= \theta^*. \tag{G12}$$

Therefore, (G12) provides a $\theta = \theta^*$ and $\lambda_2 = \lambda_2^*$ such that

$$V'(\lambda_2) + \theta^*B'(\lambda_2) = \frac{\partial}{\partial\lambda_2}[V + B] = 0$$

$$V(\lambda_2) - \theta^*gB(\lambda_2) = V - gB = 0$$

where $V'(\lambda_2) = \frac{-n\rho^{*2}}{(n+2)\lambda_2^2}$ and $B'(\lambda_2) = 2\Lambda(\lambda_2 - \frac{\rho^{*2}}{n+2})$.

Let us now substitute $V'(\lambda_2)$, $B'(\lambda_2)$, $V(\lambda_2)$ and $B(\lambda_2)$ in (G11). Then we have for (G12A)

$$- \frac{n\rho^{*2}g}{(n+2)\lambda_2^2}\{2+\frac{n\rho^{*2}}{(n+2)\lambda_2}\}^{g-1}\{\Lambda(\lambda_2-\frac{\rho^{*2}}{n+2})^2+\frac{2\rho^{*4}(n+2-\Lambda)}{(n+2)^2(n+4)}\}$$

$$+ 2\Lambda\{2+\frac{n\rho^{*2}}{(n+2)\lambda_2}\}^{g}(\lambda_2-\frac{\rho^{*2}}{n+2}) = 0, \qquad \text{(G12A)}$$

which can be written as a function of λ_2, that is,

$$\lambda_2^3+\{\frac{2\rho^{*2}(n-2)-gn\rho^{*2}}{4(n+2)}\}\lambda_2^2+\{\frac{n\rho^{*4}(g-n-2)}{2(n+2)^2}\}\lambda_2$$

$$-\frac{gn\rho^{*2}}{4\Lambda}\{\frac{\rho^{*4}(\Lambda+2)}{(n+2)^2(n+4)}\} = 0. \qquad \text{(G13)}$$

For given values of Λ and g, the expression (G13) can be used to find the minimizing value of λ_2.

In working with an expression similar to (G13) for the values of g=1,4; n=2,3,5; Λ=.2,1,2,3,4,5 and ρ^*=1, Box and Draper [3] deduce that when there is no knowledge concerning θ and Λ but that the influence of variance and bias are assumed to be about equal, the "all-bias" design (a design which is used for minimizing J when V is assumed to be equal to zero) or a design close to the "all-bias" design is best.

For our mixture problem, since we do not know the values of the parameters θ and Λ, we can propose some designs and calculate the value of J for different values assigned to θ and Λ. In other words, for given values of n, Λ, ρ^* and g, equation (G13) can be used to calculate λ_2. Then since θ

is found from (G9), we can then use equation (G3) to find the optimal or minimum value of J.

In choosing a design which will take into account contributions from both the variance and bias of $\hat{y}$, our first choice would be a design the length of whose radius is between the two extremes (that is, the radius of the "all-bias" design and the radius of the "all-variance" design). Let us choose a design matrix associated with the scaled two-level factorial where d is the radius multiplier. In this design, the points all lie on a sphere of radius $d\sqrt{n}$ and the second moment of the design is $\lambda_2 = d^2$. Since the points of the design which minimizes V alone all lie on a sphere of radius ρ^*, and the "all-bias" design has radius $(\frac{n}{n+2})^{1/2}\rho^*$, we shall consider a design which measured from the "all-bias" design is t% larger. That is, the radius of the new design is t% closer to the "all-variance" design than is the "all-bias" design and this radius is

$$\rho_t = (\frac{n}{n+2})^{1/2}\rho^* + t\{\rho^* - (\frac{n}{n+2})^{1/2}\rho^*\}$$

$$= t\rho^* + (1 - t)(\frac{n}{n+2})^{1/2}\rho^*. \qquad \text{(G14)}$$

Also, with this design, the second moment is

$$\lambda_2 = \frac{\rho_t^2}{n}. \qquad \text{(G15)}$$

With the expression (G15) for λ_2, we can find the value of t from (G14) that gives a J value (using (G3)) closest to the optimal J value calculated when λ_2 was found

from (G13). By closest, we mean a J value which is robust for various values of Λ and g. Then with this value of t, equation (G14) is used and the radius multiplier for the new design is

$$d = \frac{\rho_t}{\sqrt{n}} . \qquad \text{(G16)}$$

Thompson and Myers [22] constructed tables of J values using a formula similar to (G13) where $\rho^*=1$. The values of t, $\Theta = \frac{\Lambda}{n}$, g and k where k=n+2 are:

$t = 1.0, 0.75, 0.60, 0.50, 0.40, 0.00$

$\Theta = 0.1(0.2)0.5, 0.6(0.1)1.0$

$g = 20, 10, 6, 4, 2, 1, 0.5, 0.25$

$k = 3, 4, 6.$

For the above values, Thompson and Myers state that the J values when t = 0.60 are the most robust in terms of departures from the optimal J values for a variety of situations. Since the n in (G15) equals k-2, the tables in [22] can be used for our problem if we put $\rho^*=1$ and n=2, 4 in (G14) and (G15) for k=4, 6 in [22] respectively.

APPENDIX H.

OBTAINING THE ESTIMATES OF THE PARAMETERS IN A SECOND-DEGREE MODEL USING A CENTRAL COMPOSITE DESIGN.

In Chapter 7, Section 7.3, the form of the design matrix as well as the form of the moment matrix of a central composite design were shown. These matrices are used to obtain the vector of estimates $\hat{\underset{\sim}{\alpha}}$. In fact, the estimates $\hat{\underset{\sim}{\alpha}}$ of $\underset{\sim}{\alpha}$ in the model $\underset{\sim}{y} = \underset{\sim}{W}\,\underset{\sim}{\alpha} + \underset{\sim}{\varepsilon}$, where the matrix $\underset{\sim}{W}$ is of the form (7.3.3), are obtained by the formula

$$\hat{\underset{\sim}{\alpha}} = (\underset{\sim}{W}'\underset{\sim}{W})^{-1}\underset{\sim}{W}'\underset{\sim}{y}. \qquad \text{(H1)}$$

It was also shown in Chapter 7, that when the central composite design is used, the settings $\pm\omega$ of the axial points are dependent on the choice of whether to use the rotatability criterion or the orthogonality criterion. In some cases, a choice between the two criteria does not have to be made in that the value of ω is the same for both. In this appendix however, to arrive at the formulas for the estimates $\hat{\underset{\sim}{\alpha}}$, we shall assume that the orthogonality criterion is chosen and in addition, we let q=2.

Let us denote the observation vector $\underset{\sim}{y}$ by

$$\underset{\sim}{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_M \\ y_{M+1} \\ \vdots \\ y_{M+2p} \\ y_{M+2p+1} \\ \vdots \\ y_N \end{bmatrix} \qquad \text{(H2)}$$

where in (H2), M is the number of rows in the factorial portion of the design matrix and p (p=k-2) is the number of design variables. If the matrix $\underset{\sim}{W}$ is defined as in (7.3.3), the vector $\underset{\sim}{W}'\underset{\sim}{y}$ where $\underset{\sim}{y}$ is defined as in (H2) is

$$\underset{\sim}{W}'\underset{\sim}{y} = \begin{bmatrix} \sum_{u=1}^{N} y_u \\ c\sum_{u=1}^{M} w_{u1}y_u + \omega(y_{M+2} - y_{M+1}) \\ \vdots \\ c\sum_{u=1}^{M} w_{up}y_u + \omega(y_{M+2p} - y_{M+2p-1}) \\ c^2\sum_{u=1}^{M} y_u + \omega^2(y_{M+1} + y_{M+2}) \\ \vdots \\ c^2\sum_{u=1}^{M} y_u + \omega^2(y_{M+2p-1} + y_{M+2p}) \\ \sqrt{2}c^2\sum_{u=1}^{M} w_{u1}w_{u2}y_u \\ \vdots \\ \sqrt{2}c^2\sum_{u=1}^{M} w_{up-1}w_{up}y_u \\ L_1 \end{bmatrix} \qquad \text{(H3)}$$

where c is the radius multiplier, L_1 is a contrast of the observations ($L_1=z_{11}y_1+\ldots+z_{1N}y_N$), and $w_{ui}= \pm 1$ depending on the sign of the element in the u^{th} row and i^{th} column of the factorial portion $\underset{\sim}{D}_1$ of the matrix $\underset{\sim}{W}$.

To get the estimates $\hat{\underset{\sim}{\alpha}}$ by the equation (H1), we multiply the vector in (H3) by the matrix $(\underset{\sim}{W}'\underset{\sim}{W})^{-1} = \frac{N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}}{N}$ where the matrix $N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}$ is shown in (7.4.12). Note that the elements of the matrix $N(\underset{\sim}{W}'\underset{\sim}{W})^{-1}$ are defined in (7.4.13). Then the formulas for the estimates $\hat{\underset{\sim}{\alpha}}$ are

$$\hat{\alpha}_0 = \frac{e}{N}\sum_{u=1}^{N} y_u + \frac{f}{N}\Big[\, c^2 p \sum_{u=1}^{M} y_u + \omega^2 \sum_{u=M+1}^{M+2p} y_u\Big]$$

$$\hat{\alpha}_i = \frac{1}{N\lambda_2}\Big[\, c\sum_{u=1}^{M} w_{ui} y_u + \omega(y_{M+2i} - y_{M+2i-1})\Big] \qquad (1 \le i \le p)$$

$$\hat{\alpha}_{ii} = \frac{f}{N}\sum_{u=1}^{N} y_u + \frac{g}{N}\Big[\, c^2 \sum_{u=1}^{M} y_u + \omega^2 (y_{M+2i} + y_{M+2i-1})\Big] \qquad \text{(H4)}$$

$$+ \frac{h}{N}\Big[\, c^2 (p-1) \sum_{u=1}^{M} y_u + \omega^2 \sum_{\substack{u=M+1 \\ u\neq M+2i \\ u\neq M+2i-1}}^{M+2p} y_u\Big] \qquad (1 \le i \le p)$$

$$\frac{\hat{\alpha}_{ij}}{\sqrt{2}} = \frac{1}{2N\lambda_4}\Big[\sqrt{2}c^2 \sum_{u=1}^{M} w_{ui} w_{uj} y_u\Big] \qquad (1 \le i < j \le p)$$

$$\hat{\alpha}_{L_1} = \frac{L_1}{dN} .$$

The quantities e, f, g, and h in (H4) are linear functions of the moments of the design (see (7.4.13)). If we replace the quantities e, f, g, h in (H4) by the equivalent functions of the moments, the formulas for the estimates $\hat{\underset{\sim}{\alpha}}$ become

$$\hat{\alpha}_0 = \frac{G + \lambda_4(p-1)}{NQ}\sum_{u=1}^{N} y_u - \frac{\lambda_2}{NQ}[c^2 p\sum_{u=1}^{M} y_u + \omega^2\sum_{u=M+1}^{M+2p} y_u]$$

$$\hat{\alpha}_i = \frac{1}{N\lambda_2}[c\sum_{u=1}^{M} w_{ui}y_u + \omega(y_{M+2i} - y_{M+2i-1})] \qquad (1\le i\le p)$$

$$\hat{\alpha}_{ii} = -\frac{\lambda_2}{NQ}\sum_{u=1}^{N} y_u + \frac{1}{(G-\lambda_4)N}[c^2\sum_{u=1}^{M} y_u + \omega^2(y_{M+2i} + y_{M+2i-1})]$$

$$+ \frac{(\lambda_2^2 - \lambda_4)}{(G-\lambda_4)NQ}[c^2 p\sum_{u=1}^{M} y_u + \omega^2\sum_{u=M+1}^{M+2p} y_u] \qquad (1\le i\le p) \qquad \text{(H5)}$$

$$\hat{\alpha}_{ij} = \frac{c^2}{N\lambda_4}\sum_{u=1}^{M} w_{ui}w_{uj}y_u \qquad (1\le i<j\le p)$$

$$\hat{\alpha}_{L_1} = \frac{L_1}{dN}$$

where $Q = G - \lambda_2^2 p + \lambda_4(p-1)$. Now if we replace the moments in (H5) with design parameters (see (7.3.4)), that is,

$$\omega = Lc \quad \text{(L is explained in Table II.)}$$

$$G = \frac{1}{N}[(M+2L^4)c^4]$$

$$\lambda_4 = \frac{Mc^4}{N}$$

$$\lambda_2 = \frac{1}{N}[(M+2L^2)c^2] \quad ,$$

then the formulas in (H5) can be simplified to

$$\hat{\alpha}_0 = \frac{2L^2(L^2-p)}{D}\sum_{u=1}^{M} y_u + \frac{M(p-L^2)}{D}\sum_{u=M+1}^{M+2p} y_u + \frac{(2L^4+Mp)}{D}\sum_{u=M+2p+1}^{N} y_u$$

$$\hat{\alpha}_i = \frac{1}{(M+2L^2)c}\sum_{u=1}^{M} w_{ui}y_u + \frac{L}{(M+2L^2)c}(y_{M+2i} - y_{M+2i-1}) \quad (1\le i\le p)$$

$$\hat{\alpha}_{ii} = -\frac{(M+2L^2)}{c^2D}\sum_{u=1}^{N} y_u + \frac{1}{2L^4c^4}[c^2\sum_{u=1}^{M} y_u + L^2c^2(y_{M+2i} + y_{M+2i-1})]$$

$$+ \frac{(M+2L^2)^2}{2L^4Dc^4}[c^2p\sum_{u=1}^{M} y_u + L^2c^2\sum_{u=M+1}^{M+2p} y_u] \qquad \text{(H6)}$$

$$\hat{\alpha}_{ij} = \frac{1}{Mc^2}\sum_{u=1}^{M} w_{ui}w_{uj}y_u \qquad (1\le i<j\le p)$$

$$\hat{\alpha}_{L_1} = \frac{L_1}{dN}$$

where $D = (2L^4+Mp)N - (M+2L^2)^2p$.

Finally, when the number of independent design variables p=2, 4, 5($\frac{1}{2}$ rep.), 6 and 6($\frac{1}{4}$ rep.), the value of ω is

$$\omega = \sqrt[4]{M}\, c$$

which sets $L = \sqrt[4]{M}$, since the design is rotatable. Therefore, the formulas in (H6) are further simplified to

$$\hat{\alpha}_0 = \frac{2(1-\frac{p}{\sqrt{M}})}{D^*}\sum_{u=1}^{M} Y_u + \frac{(p-\sqrt{M})}{D^*}\sum_{u=M+1}^{M+2p} Y_u + \frac{(p+2)}{D^*}\sum_{u=M+2p+1}^{N} Y_u$$

$$\hat{\alpha}_i = \frac{1}{H^*}\sum_{u=1}^{M} w_{ui}Y_u + \frac{\sqrt[4]{M}}{H^*}(Y_{M+2i} - Y_{M+2i-1}) \qquad (1\le i\le p)$$

$$\hat{\alpha}_{ii} = \frac{(N-M-2\sqrt{M})}{c^2MD^*}\sum_{u=1}^{M} Y_u + \frac{(M+2\sqrt{M}-N)}{2\sqrt{M}c^2D^*}\sum_{u=M+1}^{M+2p} Y_u - \frac{(1+\frac{2}{\sqrt{M}})}{c^2D^*}\sum_{u=M+2p+1}^{N} Y_u$$

$$+ \frac{1}{2\sqrt{M}c^2}(Y_{M+2i} + Y_{M+2i-1}) \qquad (1\le i\le p)$$

$$\hat{\alpha}_{ij} = \frac{1}{Mc^2}\sum_{u=1}^{M} w_{ui}w_{uj}Y_u \qquad (1\le i<j\le p)$$

$$\hat{\alpha}_{L_1} = \frac{L_1}{dN}$$

where $D^* = \frac{D}{M} = N(p+2) - (2+\sqrt{M})^2 p$, $H^* = (M+2\sqrt{M})c$ and

$$L_1 = \sum_{u=1}^{N} z_{1u}Y_u \cdot$$

APPENDIX I.

MEASURING LACK OF FIT AND SUGGESTED TWO-LEVEL FACTORIALS FOR THE CASES OF TWO AND THREE CATEGORIES OF COMPONENTS.

We have failed to mention anything about the method of measuring the lack of fit of the fitted models or sizes of the designs in terms of the number of design points. In Chapters 4, 5 and 6, we talked about the fitting of a first-degree model and considered the case only where the true form of the response surface was at most second degree. In Chapter 7, we assumed that the fitting of a second-degree model was adequate. Therefore, in this appendix, we are going to discuss the method for measuring the lack of fit of the model when a first-degree model in the design variables is fitted and also list some two-level factorial designs which can be used when $q = 2, 3$.

To recapitulate, the reasons for fitting first-degree models are:

(i) the experimenter is convinced that in the particular region of interest, a linear approximation to the response surface is adequate; or

(ii) in an attempt to locate a region of maximum response through a sequence of small experiments, the simplest model is fitted initially to facilitate getting results.

In both of the above cases, it should be clear that some

measure of model inadequacy of the first-degree model must be obtained. If the experimenter has some prior measure of the experimental error variance and can be reasonably confident that this measure has not changed through time, the only concern then is in obtaining some measure of the lack of fit of the model. However, usually the experimenter does not have a measure of the error variance or one which he feels is adequate and consequently must plan to acquire, from the N experiments, a measure of this variation along with a measure of lack of fit. This is the assumption which we make here.

There are basically two methods for obtaining an estimate of the experimental error. The first method is by replicating experiments at some design points or at most, replicating all of the N experiments. When this is done, the precision of the estimates of the coefficients of the terms in the model which are affected by the replications is increased. However, the measure of the lack of fit of the model is obtained only from those terms in the model of degree higher than 1 (for example, crossproduct terms in the model sometimes used when performing a two-level factorial design).

The second method which can be used for obtaining a measure of the experimental error is to run n_0 center point replicates with each compartment of $\underset{\sim}{D}_w$ corresponding to its (horizontal) partitioning. Using this method, not only can

an estimate of error (which is based on n_0-1 degrees of freedom from each corresponding group of observations) be obtained but also we can acquire a measure of the lack of fit of the linear model. We show this with the following example. Let $p=3$, $N=2^3 + 2n_0$ and $q=2$. Write the $\underset{\sim}{W}$ matrix as

$$\underset{\sim}{W} = \left[\begin{array}{cccccccc|ccc} 1, & w_{u1}, & w_{u2}, & w_{u3}, & w_{u1}w_{u2}, & w_{u1}w_{u3}, & w_{u2}w_{u3}, & z_{1u} & w_{u1}^2, & w_{u2}^2, & w_{u3}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array}\right] \quad (I1)$$

where the values of the elements w_{ui} $(1 \le i \le 3)$ are ± 1 and 0. The columns in the left compartment of $\underset{\sim}{W}$ correspond to the terms in the fitted model and the columns in the right compartment (last three columns) correspond to the terms omitted from the model. That is, we have separated the columns corresponding to the pure quadratic terms to represent the ignored terms and at the same time included the crossproduct terms in the model.

In the analysis of variance, if we write the fitted model as

$$y_u = \alpha_0 + \sum_{i=1}^{3} \alpha_i w_{ui} + \sum\sum_{i<j}^{3} \alpha_{ij} w_{ui} w_{uj} + \alpha_{L_1} z_{1u} + \epsilon_u,$$

then the breakdown of the sums of squares would be the following:

Due to linear terms (including mean)

$$\frac{\hat{\alpha}_0}{N}\sum_{u=1}^{N} y_u + \sum_{i=1}^{3} \hat{\alpha}_i \frac{\sum_{u=1}^{N} w_{ui} y_u}{\sum_{u=1}^{N} w_{ui}^2}$$

Lack of fit (crossproduct terms)

$$\sum_{i=1}^{2}\sum_{j=2}^{3} \hat{\alpha}_{ij} \frac{\sum_{u=1}^{N} w_{ui} w_{uj} y_u}{\sum_{u=1}^{N} (w_{ui} w_{uj})^2}$$

Due to contrast term

$$\hat{\alpha}_{L_1} \frac{L_1}{\sum_{u=1}^{N} z_{1u}^2}$$

Error

$$\sum_{u=1}^{N} y_u^2 - \hat{\underline{\alpha}}' \underline{W}' \underline{y} \quad .$$

By including the crossproduct terms in the model, we have removed a source of variation from the error sum of squares and thus can use this source of variation as a measure of the lack of fit of the model.

Let us write the model now (omitting the contrast term) as

$$y_u = \alpha_0 + \sum_{i=1}^{3} \alpha_i w_{ui} + \sum\sum_{i<j}^{3} \alpha_{ij} w_{ui} w_{uj} + \sum_{i=1}^{3} \alpha_{ii} w_{ui}^2 + \epsilon_u .$$

The first and $(2^3 + 1)^{st}$ normal equations are

$$(2^3 + 2n_0)\hat{\alpha}_0 + 2^3 \left(\sum_{i=1}^{3} \hat{\alpha}_{ii} \right) = \sum_{u=1}^{2^3} y_u + \sum_{u=2^3+1}^{N} y_u \qquad (I2)$$

$$2^3 \hat{\alpha}_0 + 2^3 \left(\sum_{i=1}^{3} \hat{\alpha}_{ii} \right) = \sum_{u=1}^{2^3} y_u , \qquad (I3)$$

where $\hat{\alpha}_0$ is the estimate of the constant term in the model, $\hat{\alpha}_{ii}$ $(1 \le i \le 3)$ is the estimate of the quadratic effect of the i^{th} variable, y_u $(1 \le u \le 2^3)$ are the observations taken at the design points of the factorial design and y_u $(2^3+1 \le u \le N)$ are the observations taken at the $2n_0$ center point replicates. Subtract (I3) from (I2). We obtain

$$2n_0 \hat{\alpha}_0 = \sum_{u=2^3+1}^{N} y_u$$

and therefore,

$$\hat{\alpha}_0 = \bar{y}_{2n_0} . \qquad (I4)$$

Thus the average value of the observations taken at the $2n_0$ center point replicates provides the estimate of the mean parameter. Now, substitute (I4) in (I3). We obtain

$$2^3 \bar{y}_{2n_0} + 2^3 \left(\sum_{i=1}^{3} \hat{\alpha}_{ii} \right) = \sum_{u=1}^{2^3} y_u$$

and hence,

$$\sum_{i=1}^{3} \hat{\alpha}_{ii} = \bar{y}_{2^3} - \bar{y}_{2n_0} \cdot \qquad \text{(I5)}$$

Equation (I5) shows that the difference between the average value of the observations taken at the extreme points (vertices of the 3-dimensional cube) of the design and the average value of the observations taken at the center of the design is equal to the sum of the estimates of the coefficients of the pure quadratic terms. Therefore, we have an additional measure of the lack of fit of the model by taking center point replicates and performing the subtraction of (I5).

In the cases where q (the number of categories) is odd, we saw from the three category problem (Chapter 6) that center point replicates in $\underset{\sim}{D}_w$ were necessary in order to maintain the symmetry of the halves of $\underset{\sim}{D}_w$ as well as keep the number of observations at a minimum. When q is even, the design matrix consisted only of the settings of the design variables at the extreme points of the design. Hence, in order to obtain some measure of the experimental error for measuring the lack of fit of the first-degree model when q is even, we shall augment the design matrix $\underset{\sim}{D}_w$ with center point replicates because of the information gained as seen from equation (I5). We assumed this additional experimentation was performed when we discussed the central composite designs in Chapter 7.

The two types of first-degree designs considered in Chapter 5 were the double-simplex and the scaled two-level factorial designs. The principal advantage of the simplex designs over the scaled two-level factorial designs is the economy of design points. For practical purposes however, the use of simplex designs is likely to be limited since they supply no degrees of freedom for measuring the lack of fit and in most cases the simplex designs are less convenient to work with than are the two-level factorial designs. This was evident in Chapter 5 when we mentioned that the design matrix for a simplex design could not be generalized but that the form of the design matrix for the two-level factorial design could be generalized.

The discussion on testing lack of fit of the model has been directed toward performing this test after the first group of experiments (the top compartment of the design matrix) has been completed. In the cases where $q = 2, 3$, the simplex designs offer only one degree of freedom from the center point replicates for checking lack of fit, whereas the factorial designs offer one degree of freedom from the center point replicates and additional degrees of freedom from the terms of degree greater than 1 in the model. This latter number of degrees of freedom will of course depend on the size of the experiment (in terms of the number of design points) in the top half of $\underset{\sim}{D}_w$.

Table III. on the following page lists fractional

Table III.

Fractional Factorial Designs Suggested For Use in the Scaled Two-level Factorial Design Matrix.

Number of Categories	p	Top Half of $D_w + n_0$	"Lack of fit" degrees of freedom
2	2	$2^2 + 2$	1 + 1
2, 3	3*	$2^3 + 2$	4 + 1
	4	$2^{4-1} + 2$	3 + 1
	5**	$2^{5-1} + 3$	10 + 1
	6	$2^{6-2} + 4$	9 + 1
	7	$2^{7-2} + n_0$	24 + 1

factorial designs for various p (number of design variables). The first number in the last column corresponds to the degrees of freedom associated with the terms other than the first-degree terms in the model. The second number in the last column is the additional degree of freedom obtained by running center point replicates. The reader is referred to Davies [9] for a discussion on confounding of factorials since confounding will not be discussed here.

Finally, in Table III., the single asterik means that in the top half of $\underset{\sim}{D}_w$, one could use a 2^{3-1} design but the main effects of the variables are aliased with the 2-factor interactions. Unless the interactions are assumed negligible, both halves of $\underset{\sim}{D}_w$ would have to be performed to separate the main effects. The double asteriks mean that a 2^{5-2} design could be run in the top half of $\underset{\sim}{D}_w$ but again the 2-factor interactions are aliased with the main effects of the variables. However, if the design is to be augmented to a second-degree design, one may perform the 2^{5-2} in the top half of $\underset{\sim}{D}_w$ for the first stage of experimentation.

VITA.

John A. Cornell was born April 29, 1941 in Westerly, Rhode Island. He graduated from Venice High School in Venice, Florida in 1958. The following fall he enrolled at the University of Florida in Gainesville, Florida. He completed the requirements for a Bachelor of Science degree in Mathematics Education in December, 1962.

The author enrolled in graduate school at the University of Florida in September, 1963. He obtained a Master of Statistics degree in June, 1965. Upon graduation, he became employed by Tennessee Eastman Company, Division of Eastman Kodak, located in Kingsport, Tennessee. In September, 1966, he entered the Department of Statistics at Virginia Polytechnic Institute where he commenced working towards the Ph.D. degree.

He is a member of Phi Gamma Delta Social Fraternity, Pi Mu Epsilon Mathematics Honorary Society, The American Statistical Association and The Institute of Mathematical Statistics.

The author is married to the former Natalie R. Ragone and they have two children, John Jr. and Kenneth. Upon completion, the author and his family will reside in Gainesville, Florida, where he has accepted the position of assistant professor of Statistics in the Department of Statistics at the University of Florida.

John A. Cornell

John A. Cornell

A RESPONSE SURFACE APPROACH TO THE MIXTURE PROBLEM WHEN THE MIXTURE COMPONENTS ARE CATEGORIZED.

by

John A. Cornell

ABSTRACT

A method is developed for experiments with mixtures where the mixture components are categorized (acids, bases, etc.), and each category of components contributes a fixed proportion to the total mixture. The number of categories of mixture components is general and each category will be represented in every mixture by one or more of its member components.

The purpose of this paper is to show how standard response surface designs and polynomial models can be used for estimating the response to mixtures of the k mixture components. The experimentation is concentrated in an ellipsoidal region chosen by the experimenter, subject to the constraints placed on the components. The selection of this region, the region of interest, permits the exclusion of work in areas not of direct interest.

The transformation from a set of linearly dependent mixture components to a set of linearly independent design variables is shown. This transformation is accomplished with the use of an orthogonal matrix. Since we want the

properties of the predictor $\hat{y}$ at a point $\underset{\sim}{w}$ to be invariant to the arbitrary elements of the transformation matrix, we choose to use rotatable designs.

Frequently, there are underlying sources of variation in the experimental program whose effects can be measured by dividing the experimentation into stages, that is, blocking the observations. With the use of orthogonal contrasts of the observations, it is shown how these effects can be measured. This concept of dividing the program of experiments into stages is extended to include second-degree designs.

The radius of the largest sphere, in the metric of the design variables, that will fit inside the factor space is derived. This sphere provides an upper bound on the size of an experimental design. This is important when one desires to use a design to minimize the average variance of $\hat{y}$ only for a first-degree model. It is also shown with an example how with the use of the largest sphere, one can cover almost all combinations of the mixture components, subject to the constraints.