This research was partially supported by National Institutes of Health, Division of General Medical Sciences, Grant No. 5 R01 GM 12868-05, and by Army Research Office, Durham, Grant No. DA-ARO-D-31-124-G910.

OPTIMAL RESPONSE SURFACE TECHNIQUES USING FOURIER SERIES AND SPHERICAL HARMONICS

by

LAWRENCE L. KUPPER

Department of Statistics
University of North Carolina at Chapel Hill

Institute of Statistics Mimeo Series No. 678

APRIL 1970

ACKNOWLEDGMENTS

I gratefully acknowledge the guidance and encouragement given to me by my adviser, Professor I. M. Chakravarti, during the progress of this research.

I also wish to thank the other members of my doctoral committee, Professors R. C. Bose, N. L. Johnson, H. A. David, J. E. Grizzle, and Harry Smith, for their interest in my work. It has been a stimulating experience being associated with the members of the faculty of the Departments of Statistics and Biostatistics at the University of North Carolina.

For financial aid, I acknowledge a National Science Foundation Traineeship and support as a research assistant in the Department of Biostatistics.

I want to express my appreciation to my wife, Margie, for her patience and understanding throughout the years of my graduate study, and I also want to thank Mrs. Gay Goss for her efficient typing of this manuscript.

Finally, I owe a very special debt of gratitude to Professor William Mendenhall, Chairman of the Department of Statistics at the University of Florida, who stimulated my initial interest in Statistics and was responsible for my decision to pursue graduate studies in his department. I shall always remember his enthusiasm and dedication and his personal interest in me.

TABLE OF CONTENTS

INTRODUCTION AND SUMMARY

The general problem of response surface exploration may be described as follows: given that the form of the true functional relationship $Ey(\underset{\sim}{x}) = \eta(\underset{\sim}{x};\underset{\sim}{\beta})$ between the expected value of a random variable $y(\underset{\sim}{x})$ and a k-vector of non-stochastic factors $\underset{\sim}{x}' = (x_1, x_2, \ldots, x_k)$ is unknown, the object is to approximate as accurately as possible, within a given region of interest $\mathcal{X}$ in the k-dimensional factor space, the form of the surface $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$ by some graduating function $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$. In a given experimental situation, one might be interested in using $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ to estimate the point $\underset{\sim}{x}_0$ in $\mathcal{X}$ at which $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$ achieves a maximum (or minimum). Such a situation often occurs, for example, in the chemical industry where researchers are interested in determining those settings of the factors (e.g., pressure, temperature, concentration) which give maximum yield or highest purity of a chemical or which minimize the cost of production.

Invariably, it has been assumed that the unknown response function $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$ could be represented at every point in a given region within the factor space by a polynomial of degree d, say, in $x_1, x_2, \ldots, x_k$ of the form

$$(1) \qquad \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \beta_{11} x_1^2 + \ldots + \beta_{kk} x_k^2 + \beta_{12} x_1 x_2 + \ldots + \beta_{k-1,k} x_{k-1} x_k + \beta_{111} x_1^3 + \ldots,$$

where $\underset{\sim}{\beta}' = (\beta_0, \beta_1, \ldots, \beta_k, \beta_{11}, \ldots, \beta_{kk}, \beta_{12}, \ldots, \beta_{k-1,k}, \beta_{111}, \ldots)$ is a $1 \times \binom{k+d}{d}$ row vector of unknown parameters. The primary motivation for considering this type of expression is that it corresponds to the partial sum of a Taylor's series expansion of $\eta(\underset{\sim}{x};\beta)$ about the origin of the factor levels, the derivatives at the origin being simple multiples of the β's.

If responses are observed at a set of $\underset{\sim}{x}$'s suitably numerous and suitably placed within the region of interest, then, in general, it is possible to obtain individual estimates b_0, b_1, b_2, etc., of β_0, β_1, β_2, etc., using the method of least squares; the graduating function $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ is then the polynomial obtained by replacing the β's by the b's in (1).

There are, on the other hand, important situations where polynomial graduating functions either cannot or should not be used. For instance, let us suppose that it is desired to study an unknown response function $\eta(\underset{\sim}{x};\beta)$ on a region of interest which is the surface of a k-dimensional hypersphere of radius ρ. Many quite interesting and important situations in the natural and physical sciences can be considered to fit into this general framework. In particular, one might be interested in knowing the distribution of electric charge on a circular or spherical conductor or in predicting the rate of steady-state heat flow or the elevation at an arbitrary point on the surface of the earth. However, it has been shown (see [5], p.217) that it is not possible, if d >1, to obtain separate least squares estimates of all the parameters in (1) if each of the $\underset{\sim}{x}$'s must be chosen to satisfy the restriction $\underset{\sim}{x}'\underset{\sim}{x} = \rho^2$.

To consider yet another situation, let us suppose that we know that $\eta(\underset{\sim}{x};\beta)$ is periodic in at least one of the x_i. In this case, it is clear that an adequate approximation to $\eta(\underset{\sim}{x};\beta)$ using a polynomial would require a very large number of terms. These would oscillate in sign and the cancellation of large positive and negative values would thus give poor precision in the estimation of $\eta(x;\beta)$. Here, one would certainly prefer to use an approximating function which is itself periodic. As a simple example, if $\eta(x;\beta) = \cos 5x$, then the partial sum of a Fourier series, say $a_0 + \sum_{m=1}^{d}(a_m \cos mx + b_m \sin mx)$, would be a graduating function of few terms with smaller coefficients than those of a polynomial. Many natural phenomena are subject to periodic behavior and, in such cases, polynomial-type graduating functions would certainly be unsatisfactory.

Motivated by the above remarks, the author attempts in this dissertation to broaden the area of response surface methodology to include the use of partial sums of Fourier series and spherical harmonics as response surface graduating functions.

Chapter I serves as an introduction to the types of regression models that will be under consideration. Several important properties of these functions which are needed for later work are discussed.

In the framework of the "approximate theory" of the optimal design of experiments, a general approach to the problem of the construction of designs for response surface exploration is described in Chapter II. Various criteria are considered and some results relating certain of these criteria are given. A general theorem on the minimization of average squared bias is presented and, based on this result and some decision-theoretic notions, the new concept of the "admissibility of a response surface design" is introduced.

The results contained in the first two chapters are used in Chapter III to show that the designs proposed for the least squares fitting of the Fourier series and spherical harmonics regression models are optimal in many important respects. The valuable relationship which is found to exist between these designs and "rotatable arrangements" of points (Box and Hunter, [5]) is used to construct "exact" designs for Fourier series regression of all orders and for spherical harmonics regression of orders one and two.

In Chapter IV, techniques are developed for locating the stationary points of fitted second-order Fourier series and spherical harmonics functions and are used to construct confidence regions for the corresponding points of surfaces under study. Numerical examples are included.

CHAPTER I

FOURIER SERIES AND SPHERICAL HARMONICS MODELS AND THEIR PROPERTIES

1. The F. S. and S. H. Models

In all that follows, we shall be concerned with two regression models.

The first is the Fourier series (F.S.) model

$$(1.1) \qquad \eta_d(\phi;\underset{\sim}{\beta}_{FS}) = \sum_{n=0}^{d} U_n(\phi;\underset{\sim}{\beta}_{FS}^{(n)}),\ 0 \le \phi \le 2\pi,$$

where

$$(1.2) \qquad U_n(\phi;\underset{\sim}{\beta}_{FS}^{(n)}) = \alpha_n \cos n\phi + \beta_n \sin n\phi,\ n = 0,1,\ldots,d.$$

It will be both convenient and useful to write $U_n(\phi;\underset{\sim}{\beta}_{FS}^{(n)})$ as

$$(1.3) \qquad U_n(\phi;\underset{\sim}{\beta}_{FS}^{(n)}) = \underset{\sim}{f}'_{(n)}(\phi)\ \underset{\sim}{\beta}_{FS}^{(n)},\ n = 0,1,\ldots,d,$$

where

$$(1.4) \quad \underset{\sim}{f}_{(0)}(\phi) = 1, \underset{\sim}{\beta}_{FS}^{(0)} = \alpha_0;\ \underset{\sim}{f}'_{(n)}(\phi) = (\cos n\phi,\ \sin n\phi), \underset{\sim}{\beta}_{FS}^{(n)'} = (\alpha_n,\beta_n),$$
$$n = 1,2,\ldots,d.$$

Thus, we may write (1.1) compactly as

$$(1.5) \qquad \eta_d(\phi;\underset{\sim}{\beta}_{FS}) = \underset{\sim}{f}'(\phi)\ \underset{\sim}{\beta}_{FS},$$

where

$$(1.6) \qquad \underset{\sim}{f}'(\phi) = (\underset{\sim}{f}'_{(0)}(\phi),\ \underset{\sim}{f}'_{(1)}(\phi),\ \ldots,\ \underset{\sim}{f}'_{(d)}(\phi))$$

and

$$(1.7) \qquad \underset{\sim}{\beta}'_{FS} = (\underset{\sim}{\beta}^{(0)'}_{FS},\ \underset{\sim}{\beta}^{(1)'}_{FS},\ \ldots,\ \underset{\sim}{\beta}^{(d)'}_{FS}).$$

Note that

$$(1.8) \qquad \underset{\sim}{f}'_{(n)}(\phi)\ \underset{\sim}{f}_{(n)}(\phi) = 1 \text{ for } n = 0,1,\ldots,d,$$

and hence that

$$(1.9) \qquad \underset{\sim}{f}'(\phi)\ \underset{\sim}{f}(\phi) = (d + 1).$$

The second model that we will study is the spherical harmonics (S.H.) model

$$(1.10) \qquad \eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH}) = \textstyle\sum_{n=0}^{d} U_n(\theta,\phi;\underset{\sim}{\beta}^{(n)}_{SH}),\ 0 \le \theta \le \pi,\ 0 \le \phi \le 2\pi,$$

where

$$(1.11) \qquad U_n(\theta,\phi;\underset{\sim}{\beta}^{(n)}_{SH}) = \textstyle\sum_{m=0}^{n}(\alpha_{nm} \cos m\phi + \beta_{nm} \sin m\phi)\ P_{nm}(\cos\theta),$$

where $P_{nm}(\cos\theta)$ is the "associated Legendre function" defined as

$$(1.12) \qquad P_{nm}(\cos\theta) = \sin^m\theta\ (d^m/d \cos\theta^m)P_n(\cos\theta),$$

$P_n(\cos\theta)$ being the Legendre polynomial of degree n with argument $\cos\theta$.

As for the F. S. model, it will be worthwhile to write (1.11) as

$$(1.13) \qquad U_n(\theta,\phi;\underset{\sim}{\beta}^{(n)}_{SH}) = \underset{\sim}{f}'_{(n)}(\theta,\phi)\ \underset{\sim}{\beta}^{(n)}_{SH},\ n = 0,1,\ldots,d,$$

where

(1.14) $$\underset{\sim}{f}_{(0)}(\theta,\phi) = 1,\ \underset{\sim}{f}'_{(n)}(\theta,\phi) = \Big[P_n(\cos\theta),\ \cos\phi\ P_{n1}(\cos\theta),$$

$$\sin\phi\ P_{n1}(\cos\theta),\ \ldots,\ \cos n\phi\ P_{nn}(\cos\theta),\ \sin n\phi\ P_{nn}(\cos\theta)\Big]$$

for $n = 1,2,\ldots,d$,

and

(1.15) $\underset{\sim}{\beta}^{(0)}_{SH} = \alpha_{00}$, $\underset{\sim}{\beta}^{(n)'}_{SH} = (\alpha_{n0},\alpha_{n1},\beta_{n1},\ldots,\alpha_{nn},\beta_{nn})$ for $n = 1,2,\ldots,d$.

Thus, we may write (1.10) in the form

(1.16) $$\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH}) = \underset{\sim}{f}'(\theta,\phi)\ \underset{\sim}{\beta}_{SH},$$

where

(1.17) $$\underset{\sim}{f}'(\theta,\phi) = (\underset{\sim}{f}'_{(0)}(\theta,\phi),\ \underset{\sim}{f}'_{(1)}(\theta,\phi),\ \ldots,\ \underset{\sim}{f}'_{(d)}(\theta,\phi))$$

and

(1.18) $$\underset{\sim}{\beta}'_{SH} = (\underset{\sim}{\beta}^{(0)'}_{SH},\ \underset{\sim}{\beta}^{(1)'}_{SH},\ \ldots,\ \underset{\sim}{\beta}^{(d)'}_{SH}).$$

When written as in (1.5) and (1.16), the F. S. and S. H. models will be said to be of "order d", the vectors $\underset{\sim}{f}(\phi)$ and $\underset{\sim}{f}(\theta,\phi)$ having $(2d + 1)$ and $(d + 1)^2$ elements, respectively. The value $(2d + 1)$ is obvious, and the value $(d + 1)^2$ follows from the fact that there are $(2n + 1)$ elements in $\underset{\sim}{f}_{(n)}(\theta,\phi)$, $n = 0,1,\ldots,d$.

Remark 1.1. The domain of definition of (1.1) can be taken to be either the closed interval $[0,2\pi]$ or the circumference of a circle; in the latter case, ϕ is the angle associated with the polar coordinate representation of a point on the circle.

Remark 1.2. The domain of definition of (1.10) can be taken to be either the rectangle $\{(\theta,\phi): 0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$ or the surface of a sphere; in the latter case, θ and ϕ are the angles associated with the spherical coordinate representation of a point on the sphere.

The expression (1.1) is recognized to be the partial sum of a Fourier series and is of period 2π in ϕ. The linear transformation $\phi = 2\pi(x-a)/(b-a)$ indicates that the F. S. regression model, in addition to providing a natural way of representing a function whose domain of definition is the circumference of a circle, can also be used either to graduate an unknown response function $\eta(x;\underset{\sim}{\beta})$ on the finite interval $[a,b]$ for values of x on that interval or to graduate an unknown periodic function $\eta(x;\underset{\sim}{\beta})$ of known period $(b-a)$ for all $x \in (-\infty,\infty)$. Clearly, the F. S. model cannot be used to represent a function on the entire real line if this function is not periodic.

The expression (1.10) can be considered to be a generalization to two dimensions of (1.1); the quantity $U_n(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(n)})$ is known as a "surface spherical harmonic of degree n". Using (1.12) and the trigonometric functions for sums and differences of angles, one can easily verify that $U_n(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(n)}) = U_n(2\pi-\theta,\phi+\pi;\underset{\sim}{\beta}_{SH}^{(n)})$ for $0 \le \theta \le \pi$ and $0 \le \phi \le \pi$, and that $U_n(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(n)}) = U_n(2\pi-\theta,\phi-\pi;\underset{\sim}{\beta}_{SH}^{(n)})$ for $0 \le \theta \le \pi$ and $\pi \le \phi \le 2\pi$. Using these relations and the fact that $\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH})$ is of period 2π in both θ and ϕ, it then follows that the behavior of $\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH})$ on the entire (θ,ϕ) - plane will be known once its behavior on the rectangle $\{(\theta,\phi): 0 \le \theta \le \pi, 0 \le \phi \le 2\pi\}$ has been determined. In this light, the pair of linear transformations $\theta = \pi(x_1-a_1)/(b_1-a_1)$ and $\phi = 2\pi(x_2-a_2)/(b_2-a_2)$ indicates that the S. H. regression model, while not only providing a natural basis for representing a function defined on the surface of a sphere, could also be used either to graduate an unknown response function $\eta(x_1,x_2;\underset{\sim}{\beta})$ on the finite rectangle $\{(x_1,x_2): a_1 \le x_1 \le b_1, a_2 \le x_2 \le b_2\}$ for values of (x_1,x_2) on that rectangle or to graduate over the whole (x_1,x_2) - plane an unknown function $\eta(x_1,x_2;\underset{\sim}{\beta})$ which is of known period $2(b_1-a_1)$ in x_1 and of known period (b_2-a_2) in x_2 and which also satisfies the quite restrictive pair of relations

$$(1.19)\qquad \eta(x_1,x_2;\underset{\sim}{\beta}) = \eta(2b_1 - x_1, x_2 + \tfrac{1}{2}(b_2 - a_2);\underset{\sim}{\beta})$$

$$\text{for } a_1 \le x_1 \le b_1,\ a_2 \le x_2 \le \tfrac{1}{2}(a_2 + b_2);$$

$$\eta(x_1,x_2;\underset{\sim}{\beta}) = \eta(2b_1 - x_1, x_2 - \tfrac{1}{2}(b_2 - a_2);\underset{\sim}{\beta})$$

$$\text{for } a_1 \le x_1 \le b_1,\ \tfrac{1}{2}(a_2 + b_2) \le x_2 \le b_2.$$

In fact, the relations (1.19) are so limiting that one can feel justified in using an S. H. regression model to graduate a function $\eta(x_1,x_2;\underset{\sim}{\beta})$ on the (x_1,x_2) - plane <u>only</u> when the region of interest has finite area.

For graduating over the whole (x_1,x_2) - plane a response function $\eta(x_1,x_2;\underset{\sim}{\beta})$ which is periodic in x_1 and x_2, a possible regression model is $\eta(\phi_1,\phi_2;\underset{\sim}{\beta}_{FS}) = \eta_{d_1}(\phi_1;\underset{\sim}{\beta}_{FS})\,\eta_{d_2}(\phi_2;\underset{\sim}{\beta}_{FS})$, the direct product of two F. S. regression functions. Clearly, $\eta(\phi_1,\phi_2;\underset{\sim}{\beta}_{FS})$, which can be written as the sum of $(2d_1+1)(2d_2+1)$ distinct terms, possesses the desired periodic properties without any of the limitations suffered by the S. H. model. A little more will be said about these notions at the end of the third chapter, but let us note now that this "direct product" idea clearly generalizes to any number of dimensions.

2. Integral Properties of the F. S. and S. H. Models

Let the "weighting function" w be an element of the set $\mathcal{W}_{\mathfrak{X}}$, where $\mathcal{W}_{\mathfrak{X}} = \{w(\underset{\sim}{x}):\ w(\underset{\sim}{x})$ is a probability measure on a space $\mathfrak{X}\}$. For example, the "uniform" weighting function assigns equal weight $(\int_{\mathfrak{X}} d\underset{\sim}{x})^{-1}$ to every point $\underset{\sim}{x} \in \mathfrak{X}$.

Definition 2.1. A sequence of functions $\{f_1(\underset{\sim}{x}),\ f_2(\underset{\sim}{x}),\ \ldots\}$ is orthogonal on a space $\mathfrak{X}$ with respect to a weighting function $w \in \mathcal{W}_{\mathfrak{X}}$ if $\int_{\mathfrak{X}} f_k(\underset{\sim}{x})\ f_\ell(\underset{\sim}{x})\ dw(\underset{\sim}{x}) = 0$ whenever $k \neq \ell$. The sequence is normalized by dividing f_k by $||f_k||$, where $||f_k||^2 = \int_{\mathfrak{X}} f_k^2(\underset{\sim}{x})\ dw(\underset{\sim}{x})$.

Now, consider the "Fourier sequence"

$$(2.1)\qquad \{1,\ \cos\phi,\ \sin\phi,\ \ldots,\ \cos n\phi,\ \sin n\phi,\ \ldots\}.$$

The relations $\cos m\phi \cos n\phi = \frac{1}{2}[\cos(m+n)\phi + \cos(m-n)\phi]$, $\sin m\phi \cos n\phi = \frac{1}{2}[\sin(m+n)\phi + \sin(m-n)\phi]$, and $\sin m\phi \sin n\phi = \frac{1}{2}[\cos(m-n)\phi - \cos(m+n)\phi]$ imply that the sequence (2.1) is orthogonal with respect to the uniform weighting function $w(\phi) = \phi/2\pi$ on each of the two regions discussed in Remark 1.1, and they also show that

$$(2.2)\qquad \int_0^{2\pi} \cos^2 n\phi\ d\phi = \int_0^{2\pi} \sin^2 n\phi\ d\phi = \pi,\ n = 1,2,\ldots\ .$$

Next, consider the sequence of functions $\{1,\ x,\ x^2,\ \ldots,\ x^n,\ \ldots\}$. They are linearly independent on every finite interval of the real line. Let

$$(2.3)\qquad \{P_0(x),\ P_1(x),\ \ldots,\ P_n(x),\ \ldots\}$$

be the sequence of functions, orthogonal on the interval $[-1,1]$ with respect to the uniform weighting function $w(x) = \frac{1}{2}(1+x)$, which is obtained from $\{1, x, x^2, \ldots, x^n, \ldots\}$ by the Gram-Schmidt orthogonalization procedure. So, by construction, $P_n(x) = \sum_{j=0}^{n} C_{nj}x^j$, $C_{nn} > 0$, $n = 0,1,2,\ldots$, and

$$(2.4) \qquad \int_{-1}^{1} P_m(x)\, P_n(x)\, dx = 0 \text{ if } m \neq n.$$

The sequence (2.3) is called the "Legendre sequence" and $P_n(x)$ is the "Legendre polynomial of degree n in x". A standard expression for $P_n(x)$ is

$$(2.5) \qquad P_n(x) = \frac{(2n-1)(2n-3)\ldots1}{n!}\left[x^n - \frac{n(n-1)}{2(2n-1)}\, x^{n-2} + \frac{n(n-1)(n-2)(n-3)}{2\cdot4\cdot(2n-1)(2n-3)}\, x^{n-4} - \ldots\right],$$

where the last term is a multiple of x if n is odd and is a multiple of x^0 if n is even. Now, (2.5) can be written as

$$(2.6) \quad P_n(x) = \frac{1}{n!2^n}\frac{d^n}{dx^n}\left[x^{2n} - nx^{2n-2} + \frac{n(n-1)}{2!}\, x^{2n-4} - \ldots + (-1)^n\right],$$

and thus, from (2.6) we have

$$(2.7) \qquad P_n(x) = (2^{-n}/n!)\ (d^n/dx^n)(x^2-1)^n,$$

which is known as Rodrigues' formula. The relation

$$(2.8) \qquad \int_{-1}^{1} P_n^2(x)\, dx = 2/(2n+1)$$

follows directly by using (2.7) and performing an n-fold integration by parts.

Now, it will be shown that the "associated Legendre sequence"

$$\{P_{1m}(x),\ P_{2m}(x),\ \ldots,\ P_{nm}(x),\ \ldots\}, \tag{2.9}$$

where $P_{nm}(x)$ is given by (1.12) for $x = \cos\theta$, is an orthogonal sequence with respect to $w(x) = (1+x)/2$ on $[-1,1]$ for fixed m. (Associated Legendre functions with different values of m are, in general, not orthogonal.) To do this, we shall make use of the following recursive relation for Legendre polynomials:

$$\begin{aligned}(1-x^2)(d^{m+1}/dx^{m+1})P_n(x) - 2mx(d^m/dx^m)P_n(x)\\ + (n+m)(n-m+1)(d^{m-1}/dx^{m-1})P_n(x) = 0.\end{aligned} \tag{2.10}$$

The multiplication of (2.10) by $(1-x^2)^{m-1}$ gives

$$\begin{aligned}&(d/dx)[(1-x^2)^m\ (d^m/dx^m)P_n(x)]\\ &= -(n+m)(n-m+1)(1-x^2)^{m-1}\ (d^{m-1}/dx^{m-1})P_n(x).\end{aligned} \tag{2.11}$$

Now, from (1.12), we have

$$\begin{aligned}\int_{-1}^{1} P_{nm}(x)\ P_{n'm}(x)\ dx &= \int_{-1}^{1} (1-x^2)^m\ (d^m/dx^m)P_n(x)\ (d^m/dx^m)P_{n'}(x)\ dx\\ &= -\int_{-1}^{1} (d^{m-1}/dx^{m-1})P_{n'}(x)\ (d/dx)[(1-x^2)^m\ (d^m/dx^m)P_n(x)]\ dx,\end{aligned}$$

or, using (2.11),

$$\begin{aligned}&\int_{-1}^{1} P_{nm}(x)\ P_{n'm}(x)\ dx\\ &= (n+m)(n-m+1)\int_{-1}^{1} (1-x^2)^{m-1}\ (d^{m-1}/dx^{m-1})P_n(x)\ (d^{m-1}/dx^{m-1})P_{n'}(x)\ dx\\ &= (n+m)(n-m+1)\int_{-1}^{1} P_{n,m-1}(x)\ P_{n',m-1}(x)\ dx.\end{aligned}$$

If we apply this reduction formula m times and appeal to (2.4) and (2.8), we obtain the important result

$$(2.12)\qquad \int_{-1}^{1} P_{nm}(x)\, P_{n'm}(x)\, dx = \begin{cases} 2(n+m)!/(2n+1)(n-m)! \text{ if } n = n', \\ 0 \text{ if } n \neq n'. \end{cases}$$

Using (1.12) and (2.5), we get a useful expansion for $P_{nm}(x)$ needed in the next section:

$$(2.13)\qquad P_{nm}(x) = \frac{(2n)!}{2^n n!(n-m)!}(1-x^2)^{m/2}\Big[x^{n-m} - \frac{(n-m)(n-m-1)}{2(2n-1)}x^{n-m-2}$$

$$+ \frac{(n-m)(n-m-1)(n-m-2)(n-m-3)}{2\cdot 4(2n-1)(2n-3)}x^{n-m-4} - \dots\Big],$$

where the expression in brackets has its last term as a multiple of x if (n-m) is odd and as a multiple of x^0 if (n-m) is even.

From the integral properties we have given for the sequences (2.1) and (2.9), it follows that the "spherical harmonic sequence"

$$(2.14)\qquad \{\cos\phi\, P_{11}(x),\ \sin\phi\, P_{11}(x),\ \cos\phi\, P_{21}(x),\ \sin\phi\, P_{21}(x),\ \cos 2\phi\, P_{22}(x),$$

$$\sin 2\phi\, P_{22}(x),\ \dots,\ \cos\phi\, P_{n1}(x),\ \sin\phi\, P_{n1}(x),\ \dots,\ \cos m\phi\, P_{nm}(x),$$

$$\sin m\phi\, P_{nm}(x),\ \dots,\ \cos n\phi\, P_{nn}(x),\ \sin n\phi\, P_{nn}(x),\ \dots\}$$

is an orthogonal sequence with respect to the uniform weighting function $w(x,\phi) = (1+x)\phi/4\pi$ on the rectangle $\{(x,\phi): -1 \leq x \leq 1,\ 0 \leq \phi \leq 2\pi\}$.

The transformation $x = \cos\theta$ can now be employed to show that the validity of the expressions (2.4), (2.8) and (2.12) and the orthogonality of the sequence (2.14) are still maintained when the regions of interest are those given in Remark 1.2. It is important to note, however, that the weighting function resulting from this transformation, namely $w(\theta,\phi) = (1-\cos\theta)\phi/4\pi$, is clearly not the uniform weighting function associated with the rectangle of Remark 1.2, but is, in fact, uniform on the surface of a sphere, since $dx d\phi = -\sin\theta\, d\theta\, d\phi$ is the element of surface area of a unit sphere.

Before proceeding, let us emphasize that, with no loss in generality, we may restrict our attention to the unit circle and unit sphere. Any adjustment to be made for a radius different from unity is simply a matter of multiplying by the appropriate scale factor.

3. Polynomial Representations and Addition Theorems for the F. S. and S. H. Models

The following well-known expansions will be needed:

$$\cos n\phi = \cos^n\phi - \frac{n(n-1)}{2!}\cos^{n-2}\phi\,\sin^2\phi + \frac{n(n-1)(n-2)(n-3)}{4!}\cos^{n-4}\phi\,\sin^4\phi - \ldots, \tag{3.1}$$

and

$$\sin n\phi = n\cos^{n-1}\phi\,\sin\phi - \frac{n(n-1)(n-2)}{3!}\cos^{n-3}\phi\,\sin^3\phi + \frac{n(n-1)(n-2)(n-3)(n-4)}{5!}\cos^{n-5}\phi\,\sin^5\phi - \ldots, \tag{3.2}$$

where, in each expansion, only those terms with non-negative exponents appear.

We now give two lemmas which will be quite useful in later work.

Lemma 3.1. The expression $U_n(\phi;\underset{\sim}{\beta}_{FS}^{(n)})$ can be written as a homogeneous rational integral algebraic expression of degree n in the elements of $\underset{\sim}{x}_{FS}$, where

$$\underset{\sim}{x}'_{FS} = (\cos\phi,\ \sin\phi) = (x_1,x_2),\quad \underset{\sim}{x}'_{FS}\,\underset{\sim}{x}_{FS} = 1. \tag{3.3}$$

Proof. The result follows directly since, from (3.1), each term of $\cos n\phi$ is of the form $C_1(n,k)\cos^{n-2k}\phi \sin^{2k}\phi = C_1(n,k)\, x_1^{n-2k}\, x_2^{2k}$ where $2k < (n+1)$ and $C_1(n,k)$ is a constant depending only on n and k, and, from (3.2), each term of $\sin n\phi$ is of the form $C_2(n,\ell)\cos^{n-2\ell-1}\phi \sin^{2\ell+1}\phi = C_2(n,\ell)x_1^{n-2\ell-1}\, x_2^{2\ell+1}$, $2\ell < n$.

Analogously, we have

Lemma 3.2. The surface spherical harmonic $U_n(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(n)})$ can be written as a homogeneous rational integral algebraic expression of degree n in the elements of $\underset{\sim}{x}_{SH}$, where

(3.4) $\underset{\sim}{x}'_{SH} = (\cos\theta,\ \cos\phi\ \sin\theta,\ \sin\phi\ \sin\theta) = (x_3,x_1,x_2)$, $\underset{\sim}{x}'_{SH}\underset{\sim}{x}_{SH} = 1$.

Proof. Using (2.5) and (3.4), we see that each term of $P_n(\cos\theta)$ is of the form $C(n,j)\cos^{n-2j}\theta = C(n,j)(x_1^2 + x_2^2 + x_3^2)^j x_3^{n-2j}$, $2j < (n+1)$. From (2.13), (3.1), and (3.4), each term of $(\cos m\phi)\, P_{nm}(\cos\theta)$ is of the form $C_1(n,m,k,\ell)\cos^{m-2k}\phi\ \sin^{2k}\phi\ \sin^m\theta\ \cos^{n-2\ell-m}\theta$

$= C_1(n,m,k,\ell)(x_1^2 + x_2^2 + x_3^2)^\ell\, x_1^{m-2k}\, x_2^{2k}\, x_3^{n-2\ell-m}$, $2k < (m+1)$ and $2\ell < (n-m+1)$. And, similarly, from (2.13), (3.2), and (3.4), each term of $(\sin m\phi)\, P_{nm}(\cos\theta)$ can be put in the form

$C_2(n,m,k,\ell)(x_1^2 + x_2^2 + x_3^2)^\ell\, x_1^{m-2k-1}\, x_2^{2k+1}\, x_3^{n-2\ell-m}$, $2k < m$ and $2\ell < (n-m+1)$.

This completes the proof.

Note: The author makes no claim to originality for the results of Section 2 or for the contents of Lemmas 3.1 and 3.2. These properties of Fourier series and spherical harmonics are well known to those conversant in this subject area.

Remark 3.1. There are two very important implications contained in the above lemmas which will be quite useful in later work. The first one is that $\eta_d(\phi;\underset{\sim}{\beta}_{FS})$ is a function of ϕ only through the elements of $\underset{\sim}{x}_{FS}$ and that $\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH})$ is a function of θ and ϕ only through the elements of $\underset{\sim}{x}_{SH}$. Secondly, the expressions $\eta_d(\phi;\underset{\sim}{\beta}_{FS})$ and $\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH})$ can be written as polynomials of degree d in the elements of $\underset{\sim}{x}_{FS}$ and $\underset{\sim}{x}_{SH}$, respectively, although such representations are not unique due to the restriction $\underset{\sim}{x}'_{FS}\,\underset{\sim}{x}_{FS} = \underset{\sim}{x}'_{SH}\,\underset{\sim}{x}_{SH} = 1$. For example, it is easy to show that

$$(3.5)\quad \eta_2(\phi;\underset{\sim}{\beta}_{FS}) \equiv \eta_2(\underset{\sim}{x}_{FS};\underset{\sim}{\beta}_{FS}) = \alpha_0 + \alpha_1 x_1 + \beta_1 x_2 + \alpha_2 x_1^2 - \alpha_2 x_2^2 + 2\beta_2 x_1 x_2,$$

where

$$(3.6)\qquad \phi = \tan^{-1}(x_2/x_1),$$

and that

$$(3.7)\qquad \begin{aligned}\eta_2(\theta,\phi;\underset{\sim}{\beta}_{SH}) \equiv \eta_2(\underset{\sim}{x}_{SH};\underset{\sim}{\beta}_{SH}) &= (\alpha_{00}-\alpha_{20}/2) + \alpha_{11}x_1 + \beta_{11}x_2 \\ &+ \alpha_{10}x_3 + 3\alpha_{22}x_1^2 - 3\alpha_{22}x_2^2 + (3\alpha_{20}/2)\,x_3^2 + 6\beta_{22}x_1x_2 \\ &+ 3\alpha_{21}x_1x_3 + 3\beta_{21}x_2x_3,\end{aligned}$$

where

$$(3.8)\qquad \theta = \tan^{-1}[(x_1^2 + x_2^2)^{1/2}/x_3],\ \phi = \tan^{-1}(x_2/x_1)$$

Expressions (3.5) and (3.7) will be used in Chapter IV.

Finally, there is an important "addition theorem" relating the elements of $\underset{\sim}{f}_{(n)}(\theta,\phi)$ which will be needed for future work. The general result, the proof of which can be found in any standard reference on spherical harmonics (e.g., see Hobson [17], pp. 141-143), is as follows:

$$P_n(\cos\theta\cos\theta' + \sin\theta\sin\theta'\cos(\phi-\phi'))$$

$$= P_n(\cos\theta)\,P_n(\cos\theta') + 2\sum_{m=1}^{n}\frac{(n-m)!}{(n+m)!}\,P_{nm}(\cos\theta)\,P_{nm}(\cos\theta')\cos m(\phi-\phi').$$

Setting $\theta = \theta'$ and $\phi = \phi'$, and noting that $P_n(1) = 1$, we obtain a special form, namely

$$(3.9) \qquad P_n^2(\cos\theta) + 2\sum_{m=1}^{n}\frac{(n-m)!}{(n+m)!}\,P_{nm}^2(\cos\theta) = 1,\ 1 \leq n \leq d.$$

This expression can be regarded in the same light as the relation (1.8) for the F. S. model.

CHAPTER II

CONTRIBUTIONS TO THE THEORY OF THE OPTIMAL DESIGN OF EXPERIMENTS

4. Exact and Approximate Designs

The theory of the optimal design of experiments as initiated and principally developed by Elfving, Kiefer and Wolfowitz (e.g., see [12], [20], [22]) fits the following framework. Let $\underset{\sim}{f}'(\underset{\sim}{x}) = (f_1(\underset{\sim}{x}), f_2(\underset{\sim}{x}), \ldots, f_m(\underset{\sim}{x}))$ be a vector of m real-valued functions defined on a given space $\mathcal{X}$. In most applications, $\mathcal{X}$ is taken to be compact and the elements of $\underset{\sim}{f}(\underset{\sim}{x})$ taken to be continuous (so that suprema of certain functions are attained, insuring that optimum designs exist) and linearly independent (so that trivial redundancies are avoided). We assume that for each $\underset{\sim}{x}$ or "combination of factor levels" in $\mathcal{X}$ an experiment can be performed whose outcome is a random variable $y(\underset{\sim}{x})$, and that $\operatorname{Var} y(\underset{\sim}{x}) = \sigma^2$ for every $\underset{\sim}{x} \in \mathcal{X}$. It is further assumed that $y(\underset{\sim}{x})$ has an expected value of the explicit form

$$Ey(\underset{\sim}{x}) = \sum_{j=1}^{m} \beta_j f_j(\underset{\sim}{x}) = \underset{\sim}{f}'(\underset{\sim}{x})\underset{\sim}{\beta}, \tag{4.1}$$

and that the functions $f_1(\underset{\sim}{x}), f_2(\underset{\sim}{x}), \ldots, f_m(\underset{\sim}{x})$, called the "regression functions", are known to the experimenter while the elements of the parameter vector $\underset{\sim}{\beta}' = (\beta_1, \beta_2, \ldots, \beta_m)$ are unknowns to be estimated on the basis of a finite number N of uncorrelated observations $\{y(\underset{\sim}{x}_i)\}_{i=1}^{N}$.

An exact experimental design corresponds to a probability measure δ on $\mathcal{X}$ concentrating positive weights $w_1, w_2, \ldots, w_p$ at the distinct points $\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p$, respectively, where $w_i N = n_i$, $i = 1, 2, \ldots, p$, are integers, and where $\sum_{i=1}^{p} w_i = 1$. Clearly, the measure (design) δ specifies the points at which experiments take place, namely the $\{\underset{\sim}{x}_i\}_{i=1}^{p}$, and the number of experiments at each factor combination, namely n_i at $\underset{\sim}{x}_i$. The set $\{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \ldots, \underset{\sim}{x}_p\}$ where $w_i = \delta(\underset{\sim}{x}_i) > 0$ for every i, $i = 1, 2, \ldots, p$, and where $\sum_{i=1}^{p} w_i = 1$ is called the spectrum of the design δ, written $S(\delta)$. The theory involved in the construction of exact experimental designs will be referred to as the exact theory.

The $(m\times m)$ matrix

$$\frac{1}{N} X'X = \frac{1}{N} \sum_{i=1}^{p} n_i \, \underset{\sim}{f}(\underset{\sim}{x}_i) \, \underset{\sim}{f}'(\underset{\sim}{x}_i), \tag{4.2}$$

where $X' = (\underset{\sim}{f}(\underset{\sim}{x}_1), \underset{\sim}{f}(\underset{\sim}{x}_2), \ldots, \underset{\sim}{f}(\underset{\sim}{x}_N))$, is known as the information matrix of the exact design δ. If the parameter vector $\underset{\sim}{\beta}$ is estimated by fitting (4.1) using the method of least squares, thus obtaining a best linear unbiased estimate $\underset{\sim}{b}$, then the dispersion matrix of $\underset{\sim}{b}$ is $E[(\underset{\sim}{b}-\underset{\sim}{\beta})(\underset{\sim}{b}-\underset{\sim}{\beta})'] = (X'X)^{-1}\sigma^2$.

More generally, if $\Delta = \{\delta : \delta$ is an arbitrary probability measure on the Borel sets $\mathcal{B}$ of $\mathcal{X}$, where $\mathcal{B}$ includes all one-point sets$\}$, then, for each $\delta \in \Delta$, let us write

$$m_{ij}(\delta) = \int_{\mathcal{X}} f_i(\underset{\sim}{x}) \, f_j(\underset{\sim}{x}) \, d\delta(\underset{\sim}{x}), \tag{4.3}$$

and then define

$$M(\delta) = ((m_{ij}(\delta)))_{i,j=1}^{m}. \tag{4.4}$$

The following lemma, taken from Karlin and Studden [19], records several important properties of the matrix $M(\delta)$.

Lemma 4.1. Let $M(\delta)$ be defined as in (4.3) and (4.4). Then,

(i) for each $\delta \in \Delta$, $M(\delta)$ is positive semi-definite;

(ii) $|M(\delta)| = 0$ whenever $S(\delta)$ contains less than m points;

(iii) the family of matrices $M(\delta)$, $\delta \in \Delta$, is a convex, compact set;

(iv) for each $\delta \in \Delta$, the matrix $M(\delta)$ can be written as

$$\sum_{i=1}^{p} w_i \, \underset{\sim}{f}(\underset{\sim}{x}_i) \, \underset{\sim}{f}'(\underset{\sim}{x}_i), \text{ where } p \leq 1 + \frac{1}{2} m(m+1).$$

The most important consequence of this lemma is contained in part (iv) which permits us to restrict attention to measures concentrated on only a finite set of points when working with the class of information matrices $M(\delta)$, $\delta \in \Delta$. In particular, if $\delta \in \Delta$ corresponds to an exact design, then $NM(\delta) = X'X$.

Since the "smallness" of $M^{-1}(\delta)$ or the "largeness" of $M(\delta)$ intuitively suggests that $\underset{\sim}{b}$ is "close" to $\underset{\sim}{\beta}$, it is understandable that many optimal design criteria are based on minimizing by choice of $\delta \in \Delta$ some meaningful real functional of $M(\delta)$, say $|M^{-1}(\delta)|$ or tr $M^{-1}(\delta)$. Now, suppose that one finds a $\delta^* \in \Delta$ which minimizes some such function of $M(\delta)$. Clearly, it can happen that δ^* takes on values other than multiples of $1/N$ and thus does not correspond to an exact design. Nevertheless, it is worthwhile to consider this general approach to the construction of optimal designs, calling it the <u>approximate theory</u> and calling any measure $\delta \in \Delta$ an <u>approximate design</u>. The justification and convenience in allowing this greater generality in the choice of measures (designs) is that, in some cases, it permits one to give a complete characterization of certain optimal designs which is relevant for all N.

An optimum approximate design will give an exact design for each N which is optimum to within order N^{-1}. In some situations, the optimum approximate design will turn out to be an exact design for many and possibly for all N. We will see important examples of such phenomena in Chapter III.

5. Optimal Response Surface Designs

Let us recall from the introduction that the objective of any response surface investigation can be briefly summarized as one of approximating over some region $\mathcal{X}$ an unknown surface $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$ by some graduating function $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$. If we agree to take $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ to be the usual least-squares estimator which possesses certain well-known properties, then the design problem can be reasonably formulated as one of choosing the measure δ (i.e., the $\underset{\sim}{x}$'s and their corresponding weights) to make some meaningful function of the absolute deviation of $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ from $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$ as small as possible. One especially relevant choice for such a function is

$$E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b}) - \eta(\underset{\sim}{x};\underset{\sim}{\beta}))^2, \tag{5.1}$$

which is the expected mean-square error of $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ at $\underset{\sim}{x} \in \mathcal{X}$. It is easy to show that (5.1) can be written as the sum of two distinct error terms, namely

$$E(\hat{y}(\underset{\sim}{x};\underset{\sim}{b}) - E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}))^2, \tag{5.2}$$

the variance of $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ at $\underset{\sim}{x} \in \mathcal{X}$, and

$$(5.3) \qquad (E\hat{y}(\underset{\sim}{x};\underset{\sim}{b}) - \eta(\underset{\sim}{x};\underset{\sim}{\beta}))^2,$$

the <u>squared</u> <u>bias</u> of $\hat{y}(\underset{\sim}{x};\underset{\sim}{b})$ at $\underset{\sim}{x} \in \mathfrak{X}$.

Now, the approximate theory of the optimal design of experiments discussed in the previous section assumes that the regression function (4.1) exactly describes the true response at every point $\underset{\sim}{x} \in \mathfrak{X}$ and no consideration is given to the problem of design when the possibility exists that the incorrect model has been used. However, in reality, it is quite possible that there may be not only one, but two, sources of error, the first due to variance (sampling) error and the second due to the inadequacy of the fitted model in representing the true response $\eta(\underset{\sim}{x};\underset{\sim}{\beta})$. In a response surface setting, the latter source of error, a measure of which is given by (5.3), may far outweigh the former in importance (<u>e.g.</u>, see [2], [3], [8], [9], [10], [11]).

Motivated by these remarks, this author attempts to extend the concepts involved in the approximate theory of optimal design construction to permit the consideration of bias errors as well as variance errors. The remaining work in this chapter is involved with this problem and important applications of the results developed herein are given in Chapter III.

Now, let $e(\underset{\sim}{x},\delta)$ represent one of the three functions (5.1)-(5.3) suitably normalized with respect to N. In this framework, we give the following two definitions.

<u>Definition 5.1</u>. A design $\delta^* \in \Delta$ is said to be a <u>minimax</u> <u>design</u> if

$$(5.4) \qquad \text{MAX}_{\underset{\sim}{x}\in\mathfrak{X}}\, e(\underset{\sim}{x},\delta^*) \leq \text{MAX}_{\underset{\sim}{x}\in\mathfrak{X}}\, e(\underset{\sim}{x},\delta) \text{ for every } \delta \in \Delta,$$

where we assume $\text{MAX}_{\underset{\sim}{x}\in\mathfrak{X}}\, e(\underset{\sim}{x},\delta) < +\infty$ for some $\delta \in \Delta$.

Equivalently, δ^* is minimax if

$$\mathrm{MAX}_{\underset{\sim}{x} \in \mathcal{X}}\, e(\underset{\sim}{x},\delta^*) = \mathrm{MIN}_{\delta \in \Delta}\, \mathrm{MAX}_{\underset{\sim}{x} \in \mathcal{X}}\, e(\underset{\sim}{x},\delta).$$

Definition 5.2. A design $\delta^* \in \Delta$ is said to be a weighted error design with respect to some weighting function $w^* \in \mathcal{W}_{\mathcal{X}}$ if δ^* minimizes the quantity

$$(5.5) \qquad e(w^*,\delta) = \int_{\mathcal{X}} e(\underset{\sim}{x},\delta)\, dw^*(\underset{\sim}{x}).$$

A design $\delta^* \in \Delta$ satisfying (5.4) will be called a minimax variance (MV), a minimax bias (MB), or a minimax mean-square (MMS) design depending on whether $e(\underset{\sim}{x},\delta)$ has the form (5.2), (5.3), or (5.1), respectively. A design $\delta^* \in \Delta$ minimizing (5.5) will be referred to as a weighted variance (WV), a weighted bias (WB), or a weighted mean-square (WMS) design if $e(\underset{\sim}{x},\delta)$ has the form (5.2), (5.3), or (5.1), respectively.

The following decision-theoretic notions enable us to characterize a situation in which a weighted error design is also a minimax design.

Definition 5.3. A weighting function $w^* \in \mathcal{W}_{\mathcal{X}}$ is said to be least favorable if $\mathrm{MIN}_{\delta \in \Delta}\, e(w^*,\delta) = \mathrm{MAX}_{w \in \mathcal{W}_{\mathcal{X}}} \mathrm{MIN}_{\delta \in \Delta}\, e(w,\delta)$.

A least favorable weighting function is characterized by the following lemma.

Lemma 5.1. Suppose there exists a weighting function $w^* \in \mathcal{W}_{\mathcal{X}}$ and a weighted error design $\delta^* \in \Delta$ with respect to w^* such that $e(w^*,\delta^*) \geq \mathrm{MAX}_{\underset{\sim}{x} \in \mathcal{X}}\, e(\underset{\sim}{x},\delta^*)$. Then, δ^* is a minimax design, and the inequality above is actually an equality.

Proof. The conclusions of the lemma follow directly since, for every $\delta \in \Delta$, $\mathrm{MAX}_{\underset{\sim}{x} \in \mathcal{X}}\, e(\underset{\sim}{x},\delta) \geq e(w^*,\delta) \geq e(w^*,\delta^*) \geq \mathrm{MAX}_{\underset{\sim}{x} \in \mathcal{X}}\, e(\underset{\sim}{x},\delta^*) \geq e(w^*,\delta^*)$.

Remark 5.1. The conditions of Lemma 5.1 imply that w^* is a least favorable weighting function, since, for every $w \in \mathcal{W}_{\mathcal{X}}$, $\text{MIN}_{\delta\in\Delta}\, e(w,\delta) \leq e(w,\delta^*) \leq \text{MAX}_{\underset{\sim}{x}\in\mathcal{X}}\, e(\underset{\sim}{x},\delta^*) \leq e(w^*,\delta^*) = \text{MIN}_{\delta\in\Delta}\, e(w^*,\delta)$, so that $\text{MAX}_{w\in\mathcal{W}_{\mathcal{X}}}\text{MIN}_{\delta\in\Delta}\, e(w,\delta) = \text{MIN}_{\delta\in\Delta}\, e(w^*,\delta)$.

We now prove

Lemma 5.2. Suppose $\delta^* \in \Delta$ is a weighted error design with respect to some $w^* \in \mathcal{W}_{\mathcal{X}}$ and suppose $e(\underset{\sim}{x},\delta^*) = C < +\infty$, C a constant independent of $\underset{\sim}{x}$, for every $\underset{\sim}{x} \in \mathcal{X}$. Then, δ^* is a minimax design and $e(w^*,\delta^*) = C$.

Proof. Since $e(\underset{\sim}{x},\delta^*) = C$ for every $\underset{\sim}{x} \in \mathcal{X}$, we have $\text{MAX}_{\underset{\sim}{x}\in\mathcal{X}}\, e(\underset{\sim}{x},\delta^*) = C$. So, $e(w^*,\delta^*) = \int_{\mathcal{X}} e(\underset{\sim}{x},\delta^*)\, dw^*(\underset{\sim}{x}) = C = \text{MAX}_{\underset{\sim}{x}\in\mathcal{X}}\, e(\underset{\sim}{x},\delta^*)$. The result then follows by appealing to Lemma 5.1.

In Chapter III, an important situation will be presented where Lemma 5.2 can be put to use.

Before proceeding with the next section, let us briefly discuss some of the more important results that have been obtained in the field of optimal design of regression experiments using the approximate and exact theories. As mentioned earlier, almost all the work done in this area has involved the use of a polynomial-type model where, in the notation of Section (4.1), the $\{f_j(\underset{\sim}{x})\}_{j=1}^m$ are all the functions of the form $\Pi_{j=1}^k x_j^{r_j}$ for which the $\{r_j\}$ are nonnegative integers satisfying the relation $\sum_{j=1}^k r_j \leq d$ and where $\mathcal{X}$ is a subset of $\mathcal{R}^k$, Euclidean k-space. Note that, in this case, $m = \binom{k+d}{d}$.

Some elegant results have been obtained in the framework of the approximate theory when the regression models are polynomials (see [19] for a general treatise on the contributions of various authors in this area). For example, it has been shown for polynomial regression of order d that, when k = 1, and $\mathcal{X} = [-1,1]$, $\text{MAX}_{x\epsilon[-1,1]}\ \underset{\sim}{f}'(x)M^{-1}(\delta)\underset{\sim}{f}(x)$ is minimized or, by the Equivalence Theorem of Kiefer and Wolfowitz [23], $|M(\delta)|$ is maximized by that $\delta^* \epsilon \Lambda$ concentrating equal weights $1/(d+1)$ at the zeros of the polynomial $(1-x^2)\ P_d'(x)$. (Designs satisfying the first criterion above are said to be G-optimal and those which satisfy the second are D-optimal). Kiefer, in [21], proves a general result which shows when $\mathcal{X} = \{\underset{\sim}{x}:\underset{\sim}{x}'\underset{\sim}{x} \leq 1\}$ (the unit k-sphere) that some but not all of the class of designs known as rotatable designs, which were originally proposed by Box and Hunter [5] and then considered by others (e.g., [1], [14]), are D- and G-optimal. In the framework of the approximate theory, this work by Kiefer represents the only attempt that has been made in the literature to justify the intuitive appeal of these designs. This author gives a somewhat related result in Section 9 which characterizes rotatable arrangements of points as optimal designs for F. S. and S. H. regression.

Using the exact theory, several authors have considered the problem of the construction of optimal regression designs when the degree d_1 of the least squares polynomial is exactly one less than the degree d_2 of the polynomial representing the true response. For example, both David and Arens [7] and Box and Draper [2] consider the problem of constructing WMS designs with respect to the uniform weighting function on $[-1,1]$ when $d_1 = 1$, $d_2 = 2$. The former pair of authors also construct MMS designs in the above setting and are the only ones to do so, mainly because, for general d and k, the construction of MB and MMS designs for polynomial regression appears to be quite a formidable task. The latter pair investigate the choice of designs for polynomial regression on the unit k-sphere, when $d_1 = 1$, $d_2 = 2$ in [2] and when $d_1 = 2$,

$d_2 = 3$ in [3], by first finding a general class of WMS designs and then selecting from these a subclass for which the non-centrality term in the expectation of the residual sum of squares is large. Draper and Lawrence, in [8], [9] and [10], find WB designs when $d_1 = 1$, $d_2 = 2$, and $d_1 = 2$, $d_2 = 3$ where $\mathcal{X}$ is the unit k-cube in [8], the three-dimensional simplex (equilateral triangle) in [9], and the four-dimensional simplex (tetrahedron) in [10], and, as do the authors in [2] and [3], they show that the designs appropriate when both variance and bias occur are not very different from these "all-bias" designs. In [11], sequential second-order WB designs are constructed so that both the first-order portion and the second-order portion give protection against biases of one order higher. Here, rather than a uniform weighting function, the authors use a family of symmetric multivariate distribution weight functions depending on a single parameter to find a series of optimal designs for polynomial regression.

6. A General Result on the Construction of Weighted Bias Designs

Let $\hat{y}(\underset{\sim}{x};\underset{\sim}{b}_1) = \underset{\sim}{f}'_{(1)}(\underset{\sim}{x})\underset{\sim}{b}_1$ and $Ey(\underset{\sim}{x}) = \eta(\underset{\sim}{x};\underset{\sim}{\beta}_1,\underset{\sim}{\beta}_2) = \underset{\sim}{f}'_{(1)}(\underset{\sim}{x})\underset{\sim}{\beta}_1 + \underset{\sim}{f}'_{(2)}(\underset{\sim}{x})\underset{\sim}{\beta}_2$ denote the forms of the estimated and true responses at $\underset{\sim}{x} \in \mathcal{X}$, respectively; the vectors $\underset{\sim}{f}'_{(1)}(\underset{\sim}{x}) = (f_1(\underset{\sim}{x}), f_2(\underset{\sim}{x}), \ldots, f_{m_1}(\underset{\sim}{x}))$ and $\underset{\sim}{f}'_{(2)}(\underset{\sim}{x}) = (f_{m_1+1}(\underset{\sim}{x}), f_{m_1+2}(\underset{\sim}{x}), \ldots, f_{m_2}(\underset{\sim}{x}))$ are assumed to be known while the parameter vectors $\underset{\sim}{\beta}'_1 = (\beta_1, \beta_2, \ldots, \beta_{m_1})$ and $\underset{\sim}{\beta}'_2 = (\beta_{m_1+1}, \beta_{m_1+2}, \ldots, \beta_{m_2})$ are unknown, where $m_2 > m_1 \geq 1$.

With observations at $N = \sum_{i=1}^{p} n_i$ points in $\mathcal{X}$, we define the matrices $X'_1 = (\underset{\sim}{f}_{(1)}(\underset{\sim}{x}_1), \underset{\sim}{f}_{(1)}(\underset{\sim}{x}_2), \ldots, \underset{\sim}{f}_{(1)}(\underset{\sim}{x}_N))$ and $X'_2 = (\underset{\sim}{f}_{(2)}(\underset{\sim}{x}_1), \underset{\sim}{f}_{(2)}(\underset{\sim}{x}_2), \ldots, \underset{\sim}{f}_{(2)}(\underset{\sim}{x}_N))$, and we further assume that X_1 is of full rank. Then, if we agree to take b_1 as the standard least squares estimator of the form $(X_1'X_1)^{-1}X_1'y$, where $y' = (y(x_1), y(x_2), \ldots, y(x_N))$ is the vector of observations, it follows that $E\hat{y}(x;b_1) = f'_{(1)}(x)\beta_1 + f'_{(1)}(x)A\beta_2$, the matrix $A = (X_1'X_1)^{-1} X_1'X_2$ being commonly referred to as the "alias matrix", where $X_1'X_1 = \sum_{i=1}^{p} n_i f_{(1)}(x_i)\, f'_{(1)}(x_i)$ and $X_1'X_2 = \sum_{i=1}^{p} n_i f_{(1)}(x_i)\, f'_{(2)}(x_i)$. (Note that if a_i' is the i-th row of A, then $Eb_i = \beta_i + a_i'\beta_2$, $i = 1, 2, \ldots, m_1$; in particular, if $A = 0$, then $Eb_1 = \beta_1$.).

So, from these results, it is clear that we can write the "squared bias" $(E\hat{y}(x;b_1) - \eta(x;\beta_1,\beta_2))^2$ as $[(f'_{(1)}(x)\ A - f'_{(2)}(x))\beta_2]^2$, which is a function of the choice of the exact design δ only through the matrix A.

Now, since A can be written in the suggestive form $N(X_1'X_1)^{-1} \frac{1}{N} X_1'X_2$, it is meaningful to work in the more general setting of the approximate theory and to consider minimizing the integrated squared bias $B_\delta(w;\beta_2) = \int_{\mathcal{X}} B_\delta(x;\beta_2)\, dw(x)$ by choice of $\delta \in \Delta$, where $w \in \mathcal{W}_{\mathcal{X}}$ and where

(6.1) $$B_\delta(x;\beta_2) = [(f'_{(1)}(x)\ A(\delta) - f'_{(2)}(x))\beta_2]^2,$$

(6.2) $$A(\delta) = M_{11}^{-1}(\delta)\ M_{12}(\delta),$$

(6.3) $$M_{11}(\delta) = ((m_{ij}(\delta)))_{i,j=1}^{m_1},$$

and

(6.4) $$M_{12}(\delta) = ((m_{ij}(\delta)))_{i=1,\, j=m_1+1}^{m_1,\ m_2},$$

the general expression for $m_{ij}(\delta)$ being given by (4.3). Note that we are taking $M_{11}(\delta)$ to be non-singular.

So, proceeding in this way, we have

$$B_\delta(w;\underset{\sim}{\beta}_2) = \int_{\mathcal{X}} \underset{\sim}{\beta}_2'[A'(\delta)\ \underset{\sim}{f}_{(1)}(\underset{\sim}{x}) - \underset{\sim}{f}_{(2)}(\underset{\sim}{x})][\underset{\sim}{f}_{(1)}'(\underset{\sim}{x})\ A(\delta) - \underset{\sim}{f}_{(2)}'(\underset{\sim}{x})]\underset{\sim}{\beta}_2\ dw(\underset{\sim}{x})$$

$$= \underset{\sim}{\beta}_2'\ \Gamma\ \underset{\sim}{\beta}_2, \text{ where } \Gamma = A'(\delta)\Lambda_{11}A(\delta) - \Lambda_{12}'A(\delta) - A'(\delta)\Lambda_{12} + \Lambda_{22},$$

with $\Lambda_{11} = \int_{\mathcal{X}} \underset{\sim}{f}_{(1)}(\underset{\sim}{x})\ \underset{\sim}{f}_{(1)}'(\underset{\sim}{x})\ dw(\underset{\sim}{x})$, $\Lambda_{12} = \int_{\mathcal{X}} \underset{\sim}{f}_{(1)}(\underset{\sim}{x})\ \underset{\sim}{f}_{(2)}'(\underset{\sim}{x})\ dw(\underset{\sim}{x})$,

and $\Lambda_{22} = \int_{\mathcal{X}} \underset{\sim}{f}_{(2)}(\underset{\sim}{x})\ \underset{\sim}{f}_{(2)}'(\underset{\sim}{x})\ dw(\underset{\sim}{x})$. Now, assuming that Λ_{11} is non-singular, we can write $\Gamma = (\Lambda_{22} - \Lambda_{12}'\ \Lambda_{11}^{-1}\ \Lambda_{12})$

$$+ (A(\delta) - \Lambda_{11}^{-1}\Lambda_{12})'\ \Lambda_{11}(A(\delta) - \Lambda_{11}^{-1}\Lambda_{12})$$

$$= \Gamma_1 + \Gamma_2.$$

Since $0 \le \int_{\mathcal{X}} [(\underset{\sim}{f}_{(1)}'(\underset{\sim}{x}),\ \underset{\sim}{f}_{(2)}'(\underset{\sim}{x}))\underset{\sim}{\alpha}]^2\ dw(\underset{\sim}{x}) = \underset{\sim}{\alpha}' \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix} \underset{\sim}{\alpha}$, it follows that Λ_{11} and $\begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix}$ are both positive semi-definite.

Finally, $\begin{bmatrix} \Lambda_{11} & 0 \\ 0 & \Gamma_1 \end{bmatrix} = \begin{bmatrix} I & 0 \\ -\Lambda_{12}'\ \Lambda_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix} \begin{bmatrix} I & -\Lambda_{11}^{-1}\ \Lambda_{12} \\ 0 & I \end{bmatrix}$

$= T' \begin{bmatrix} \Lambda_{11} & \Lambda_{12} \\ \Lambda_{12}' & \Lambda_{22} \end{bmatrix} T$, which shows that Γ_1 is also positive semi-definite.

Thus, we have the following general result: no matter what the value of $\underset{\sim}{\beta}_2$, $B_\delta(w;\underset{\sim}{\beta}_2)$ is minimized by choice of $\delta \in \Delta$ when $A(\delta) = \Lambda_{11}^{-1}\Lambda_{12}$, and, in particular, when $M_{11}(\delta) = \Lambda_{11}$ and $M_{12}(\delta) = \Lambda_{12}$. In fact, from the structures of the pairs of matrices $M_{11}(\delta)$, Λ_{11} and $M_{12}(\delta)$, Λ_{12}, it follows directly that the latter two matrix equalities above can be satisfied by choice of $\delta \varepsilon \Delta$ if δ is taken to be the weighting function w itself. Since, in general, w is not discrete (e.g., it is chosen to be uniform on $\mathcal{X}$ in many instances), it is necessary to find a measure concentrated only on a finite set of points in $\mathcal{X}$ (see part (iv) of Lemma 4.1) which is equivalent to w in the sense that the pair of matrices Λ_{11} and Λ_{12} are identical for both w and its discrete counterpart.

This result follows as a direct generalization of work by Box and Draper in [2], who restrict their attention to the construction of exact WB designs where the fitted and true models are polynomials and the associated weighting function is always uniform on the given region of interest $\mathcal{X}$. The new treatment given here allows one to consider more general types of regression functions for both the fitted and true linear models and enables one to characterize the WB design with respect to any given $w \varepsilon \mathcal{W}_{\mathcal{X}}$ as an optimal approximate design. That such a generalization is meaningful and necessary will become apparent both in the next section and in the next chapter.

Motivated by the Box-Draper and Draper-Lawrence results discussed in Section 5 which emphasize the dominating influence of bias considerations in the construction of WMS designs for polynomial regression, Hader, Karson and Manson in [15] consider using a method of estimation of the coefficients of the fitted polynomial which is aimed directly at minimizing integrated squared bias with respect to a uniform weighting function, rather than assuming, as we have done, that the coefficients must be estimated by the method of least squares and that the integrated squared bias is then minimized by choice of design δ. Since a WB design is obtained by choice of estimator using this new approach, the authors can and do use the additional flexibility that remains in the choice of the location of the design points in $\mathcal{X}$ to construct WV designs. Using this technique, they are able to find designs with smaller integrated mean-square error than corresponding WMS designs that assume least squares estimation of the coefficients in the fitted polynomial.

In Chapter III, it will be shown that the class of optimal designs proposed for the least squares fitting of the F. S. and S. H. regression models minimize integrated squared bias and integrated variance simultaneously and thus no "better" class of designs could be found by using the Hader-Karson-Manson "minimum bias estimation" technique.

7. The Admissibility of a Response Surface Design

The approximate theory approach to the optimal design of experiments as described in Section 4 characterizes an optimal design $\delta^* \varepsilon \Delta$ as one which minimizes some appropriate real function of the information matrix $M(\delta)$ defined by the relations (4.3) and (4.4). In this light, the standard concepts that have so far been developed concerning the admissibility of an experimental design (see [19], pp. 808-812) have been based on the following definition.

Definition 7.1. A design $\delta \varepsilon \Delta$ is said to be admissible if there does not exist a design $\delta' \varepsilon \Delta$ such that $M(\delta') \geq M(\delta)$, where this inequality signifies that the matrix $M(\delta') - M(\delta)$ is positive semi-definite and $M(\delta') \neq M(\delta)$.

It is important to realize that this notion of admissibility is based solely on variance considerations and implicit in Definition 7.1

is the assumption that the model (4.1) exactly represents the true response at every point $\underset{\sim}{x} \in \mathfrak{X}$; in other words, no allowance is made for the possible occurrence of bias errors resulting from the use of an incorrect model. However, such as assumption is not entirely realistic in a response surface setting where a graduating function such as a polynomial will always fail, at least to some extent, to describe a response surface. The need to consider an alternative to the approximate theory concept of admissibility as expressed in Definition 7.1 is even further emphasized by results mentioned earlier ([2], [3], [8], [9], [10], [11]) which illustrate in typical situations the overriding importance of bias considerations when selecting a response surface design to minimize integrated mean-square error.

Motivated by these considerations, we now proceed to define the new concepts of V-, B-, and MS- admissibility, which have particular appeal in a response surface framework. In the notation of Section 6, let $e(\underset{\sim}{x},\delta)$ denote any of the following three functions:

(7.1) $$V_\delta(\underset{\sim}{x}) = \underset{\sim}{f}'_{(1)}(\underset{\sim}{x})\, M_{11}^{-1}(\delta)\, \underset{\sim}{f}_{(1)}(\underset{\sim}{x});$$

(7.2) $$B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2);$$

and,

(7.3) $$MS_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2) = V_\delta(\underset{\sim}{x}) + B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2),$$

where $B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, $A(\delta)$, $M_{11}(\delta)$ and $M_{12}(\delta)$ are given by (6.1)-(6.4).

Then, we have

Definition 7.2. A design $\delta \in \Delta$ is said to be variance (V-), bias (B-), or mean-square (MS-) admissible, depending on whether $e(\underset{\sim}{x},\delta)$ is equal to $V_\delta(\underset{\sim}{x})$, $B_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, or $MS_\delta(\underset{\sim}{x};\underset{\sim}{\beta}_2)$, respectively, if there does not exist a design $\delta' \in \Delta$ such that $e(\underset{\sim}{x},\delta') \leq e(\underset{\sim}{x},\delta)$ for every $\underset{\sim}{x} \in \mathcal{X}$ and $e(\underset{\sim}{x}_0,\delta') < e(\underset{\sim}{x}_0,\delta)$ for some $\underset{\sim}{x}_0 \in \mathcal{X}$.

It is easy to see that the properties of V-, B-, and MS- admissibility are certainly desirable ones for a response surface design to possess. In fact, one should feel somewhat disappointed in learning that a response surface design which minimized some meaningful function like maximum or integrated mean-square error, say, was not MS-admissible. Unfortunately, such a situation is not out of the realm of possibility; for example, a design which minimizes maximum squared bias is not necessarily B- admissible.

The following theorem, a take-off on a result in [13], p. 62, attempts to shed some light on this problem by characterizing an important situation in which a response surface design can be said to possess one of the types of admissibility introduced in Definition 7.2.

First, a preliminary definition.

Definition 7.3. Let $\mathcal{X}$ be a metric space with metric d. A point $\underset{\sim}{x}_0 \in \mathcal{X}$ is said to be in the support of a weighting function $w \in \mathcal{W}_{\mathcal{X}}$ if, for every $\varepsilon > 0$, the neighborhood $N_\varepsilon(\underset{\sim}{x}_0) = \{\underset{\sim}{x} \in \mathcal{X}: d(\underset{\sim}{x},\underset{\sim}{x}_0) < \varepsilon\}$ has "positive weight" in the sense that $\int_{N_\varepsilon(\underset{\sim}{x}_0)} dw(\underset{\sim}{x}) > 0$.

Then, we have

THEOREM 7.1. Let $\mathcal{X}$ be a metric space with metric d and assume that the functions (7.1), (7.2), and (7.3) are each continuous in $\underset{\sim}{x}$ for all $\delta \in \Delta$. Then, a design $\delta^* \in \Delta$ is V-, B-, or MS- admissible depending on whether $e(\underset{\sim}{x},\delta)$ has the form (7.1), (7.2), or (7.3), respectively, if δ^* minimizes $e(w^*,\delta) = \int_{\mathcal{X}} e(\underset{\sim}{x},\delta)\, dw^*(\underset{\sim}{x})$ with respect to some $w^* \in \mathcal{W}_{\mathcal{X}}$

and if the support of w^* is $\mathcal{X}$ itself.

Proof. Assume that δ^* is not admissible in the sense of Definition 7.2. Then, there must exist a $\delta' \in \Delta$ for which $e(\underset{\sim}{x},\delta') \leq e(\underset{\sim}{x},\delta^*)$ for all $\underset{\sim}{x} \in \mathcal{X}$ and for which $\eta = e(\underset{\sim}{x}_0,\delta^*) - e(\underset{\sim}{x}_0,\delta') > 0$ for some $\underset{\sim}{x}_0 \in \mathcal{X}$. Because $e(\underset{\sim}{x},\delta)$ is continuous in $\underset{\sim}{x}$ for each $\delta \in \Delta$, there exists an $\varepsilon > 0$ such that $|e(\underset{\sim}{x},\delta^*) - e(\underset{\sim}{x}_0,\delta^*)| \leq \eta/4$ and $|e(\underset{\sim}{x},\delta') - e(\underset{\sim}{x}_0,\delta')| \leq \eta/4$ if $\underset{\sim}{x} \in N_\varepsilon(\underset{\sim}{x}_0)$. Thus, for $\underset{\sim}{x} \in N_\varepsilon(\underset{\sim}{x}_0)$, $e(\underset{\sim}{x},\delta^*) - e(\underset{\sim}{x},\delta') \geq (e(\underset{\sim}{x}_0,\delta^*) - \eta/4) - (e(\underset{\sim}{x}_0,\delta') + \eta/4) = e(\underset{\sim}{x}_0,\delta^*) - e(\underset{\sim}{x}_0,\delta') - \eta/2 = \eta/2 > 0$. Hence, we have $e(w^*,\delta^*) - e(w^*,\delta') = \int_{\mathcal{X}} (e(\underset{\sim}{x},\delta^*) - e(\underset{\sim}{x},\delta'))\, dw^*(\underset{\sim}{x}) \geq \int_{N_\varepsilon(\underset{\sim}{x}_0)} (e(\underset{\sim}{x},\delta^*) - e(\underset{\sim}{x},\delta'))\, dw^*(\underset{\sim}{x}) \geq \frac{\eta}{2} \int_{N_\varepsilon(\underset{\sim}{x}_0)} dw^*(\underset{\sim}{x})$. But, since $\underset{\sim}{x}_0$ is in the support of w^*, the last expression above is strictly greater than zero. This contradicts the assumption that δ^* minimizes $e(w^*,\delta)$ and completes the proof.

It is clear from the structure of the functions (7.1)-(7.3) that the continuity assumptions in Theorem 7.1 are not the least bit restrictive, and, in fact, obviously hold for the standard regression models such as polynomials and various trigonometric functions like (1.1) and (1.10).

Theorem 3.1 provides great incentive for constructing and using response surface designs which minimize a given measure of error between the fitted and true models that is averaged rather than, say, maximized over $\mathcal{X}$; in particular, it lends support to the work of Box and Draper and Draper and Lawrence and shows that their WV, WB, and WMS designs are V-, B-, and MS- admissible, respectively. (The designs that we will propose for F. S. and S. H. regression will be <u>simultaneously</u> V-, B-, and MS- admissible).

Finally, it is important to mention the fact that admissibility in the sense of Definition 7.1 is "stronger" than V- admissibility since a design $\delta \in \Lambda$ which is admissible in the former sense is clearly V- admissible, while the converse is not necessarily true.

CHAPTER III

OPTIMAL DESIGNS FOR F. S. AND S. H. REGRESSION

8. Optimal Properties of $M(\delta^*_{FS})$ and $M(\delta^*_{SH})$

Working in the framework of the approximate theory and using the results in Chapters I and II, we will, in this section, characterize the optimal designs δ^*_{FS} and δ^*_{SH} for F. S. and S. H. regression, respectively, of order d in terms of their corresponding information matrices $M(\delta^*_{FS})$ and $M(\delta^*_{SH})$. The many optimal properties possessed by the approximate designs δ^*_{FS} and δ^*_{SH} will be emphasized. The important question of whether δ^*_{FS} and δ^*_{SH} correspond to exact designs in certain cases will be answered in the next section.

Now, for the S. H. model, we define from (1.14) the quantities

$$(8.1)\qquad \underset{\sim}{f}_{[0]}(\theta,\phi) = \underset{\sim}{f}_{(0)}(\theta,\phi) = 1;\ \underset{\sim}{f}'_{[n]}(\theta,\phi) = \Bigg[P_n(\cos\theta),$$

$$\left(\frac{2}{n(n+1)}\right)^{1/2}\cos\phi\ P_{n1}(\cos\theta),\ \left(\frac{2}{n(n+1)}\right)^{1/2}\sin\phi\ P_{n1}(\cos\theta),\ \ldots,$$

$$\left(\frac{2(n-m)!}{(n+m)!}\right)^{1/2}\cos m\phi\ P_{nm}(\cos\theta),\ \left(\frac{2(n-m)!}{(n+m)!}\right)^{1/2}\sin m\phi\ P_{nm}(\cos\theta),$$

$$\ldots,\ \left(\frac{2}{(2n)!}\right)^{1/2}\cos n\phi\ P_{nn}(\cos\theta),\ \left(\frac{2}{(2n)!}\right)^{1/2}\sin n\phi\ P_{nn}(\cos\theta)\Bigg],$$

$n = 1, 2, \ldots, d.$

Correspondingly, we let $\underset{\sim}{\beta}_{SH}^{[n]}$ be the vector of elements of $\underset{\sim}{\beta}_{SH}^{(n)}$ with suitable multipliers attached so that $U_n(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(n)}) = \underset{\sim}{f}'_{[n]}(\theta,\phi)\underset{\sim}{\beta}_{SH}^{[n]}$; in particular, it follows from (1.15) that

$$(8.2)\quad \underset{\sim}{\beta}_{SH}^{[0]} = \underset{\sim}{\beta}_{SH}^{(0)} = \alpha_{00};\ \underset{\sim}{\beta}_{SH}^{[n]'} = \left[\alpha_{n0}, \left(\frac{n(n+1)}{2}\right)^{1/2}\alpha_{n1}, \left(\frac{n(n+1)}{2}\right)^{1/2}\beta_{n1},\right.$$
$$\ldots, \left(\frac{(n+m)!}{2(n-m)!}\right)^{1/2}\alpha_{nm}, \left(\frac{(n+m)!}{2(n-m)!}\right)^{1/2}\beta_{nm}, \ldots, \left(\frac{(2n)!}{2}\right)^{1/2}\alpha_{nn},$$
$$\left.\left(\frac{(2n)!}{2}\right)^{1/2}\beta_{nn}\right], n = 1, 2, \ldots, d.$$

So, using (8.1) and (8.2), we are able to write (1.10) as

$$(8.3)\qquad \eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH}) = \underset{\sim}{f}^{*\prime}(\theta,\phi)\ \underset{\sim}{\beta}_{SH}^{*},$$

where

$$(8.4)\qquad \underset{\sim}{f}^{*\prime}(\theta,\phi) = \left(\underset{\sim}{f}'_{[0]}(\theta,\phi), \underset{\sim}{f}'_{[1]}(\theta,\phi), \ldots, \underset{\sim}{f}'_{[d]}(\theta,\phi)\right)$$

and

$$(8.5)\qquad \underset{\sim}{\beta}_{SH}^{*\prime} = \left(\underset{\sim}{\beta}_{SH}^{[0]'}, \underset{\sim}{\beta}_{SH}^{[1]'}, \ldots, \underset{\sim}{\beta}_{SH}^{[d]'}\right).$$

It follows from the addition theorem (3.9) that

$$(8.6)\qquad \underset{\sim}{f}'_{[n]}(\theta,\phi)\ \underset{\sim}{f}_{[n]}(\theta,\phi) = 1, n = 0, 1, \ldots, d,$$

and hence that

$$(8.7)\qquad \underset{\sim}{f}^{*\prime}(\theta,\phi)\ \underset{\sim}{f}^{*}(\theta,\phi) = (d+1).$$

The relations (8.6) and (8.7) are analogous to (1.8) and (1.9) for the F. S. regression model.

First, consider the problem of finding WV designs for F. S. and S. H. regression of order d. If the model to be fitted by least squares is of the general form (4.1), then, in the framework of the approximate theory, it is desired to find a measure $\delta^* \varepsilon \Lambda$ which minimizes the quantity $\frac{N}{\sigma^2} V_\delta(w) = \int_{\mathfrak{X}} \underset{\sim}{f}'(\underset{\sim}{x}) M^{-1}(\delta) \underset{\sim}{f}(\underset{\sim}{x})\, dw(\underset{\sim}{x})$, where $M(\delta) = \sum_{i=1}^{p} w_i \underset{\sim}{f}(\underset{\sim}{x}_i) \underset{\sim}{f}'(\underset{\sim}{x}_i)$ and $w \varepsilon \mathcal{W}_{\mathfrak{X}}$. In what follows, we shall find it convenient to write $\frac{N}{\sigma^2} V_\delta(w)$ as $\frac{N}{\sigma^2} V_\delta(w) = tr[M^{-1}(\delta) \int_{\mathfrak{X}} \underset{\sim}{f}(\underset{\sim}{x})\, \underset{\sim}{f}'(\underset{\sim}{x})\, dw(\underset{\sim}{x})]$.

Now, for S. H. regression of order d, we wish to find a design δ^*_{SH} which minimizes the quantity

$$\frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH}) = \frac{1}{4\pi}\, tr\left[M^{*-1}(\delta_{SH}) \int_0^{2\pi} \int_0^{\pi} \underset{\sim}{f}^*(\theta,\phi)\, \underset{\sim}{f}^{*\prime}(\theta,\phi) \sin\theta\, d\theta d\phi\right],$$

where

$$M^*(\delta_{SH}) = \sum_{i=1}^{p} w_i\, \underset{\sim}{f}^*(\theta_i,\phi_i)\, \underset{\sim}{f}^{*\prime}(\theta_i,\phi_i) \tag{8.8}$$

and where the integration can be considered to be performed either with respect to the uniform weighting function on the surface of a sphere or with respect to the weighting function $w^*_{SH}(\theta,\phi) = (1-\cos\theta)\phi/4\pi$ on the rectangle $\{(\theta,\phi):\ 0 \le \theta \le \pi,\ 0 \le \phi \le 2\pi\}$.

<u>Note</u>: The fact that we are working with $M^*(\delta_{SH})$ instead of $M(\delta_{SH}) = \sum_{i=1}^{p} w_i\, \underset{\sim}{f}(\theta_i,\phi_i)\, \underset{\sim}{f}'(\theta_i,\phi_i)$ will not affect the characterization of the WV design δ^*_{SH}. More will be said about this later in the section.

From (8.4), it is easy to see that $\underset{\sim}{f}^*(\theta,\phi)\, \underset{\sim}{f}^{*\prime}(\theta,\phi)$ is a $(d+1)^2 \times (d+1)^2$ block matrix with its typical element being the $(2n+1) \times (2n'+1)$ matrix $\underset{\sim}{f}_{[n]}(\theta,\phi)\, \underset{\sim}{f}'_{[n']}(\theta,\phi)$, $n = 0, 1, \ldots, d$, $n' = 0, 1, \ldots, d$. Using (2.2), (2.8), (2.12) and the orthogonality properties of the "spherical harmonic sequence" discussed in Section 2, we can write

$$(8.9)\qquad \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} \underset{\sim}{f}_{[n]}(\theta,\phi)\ \underset{\sim}{f}'_{[n']}(\theta,\phi)\ \sin\theta\ d\theta d\phi = \begin{cases} 0, & n\neq n' \\ \frac{1}{(2n+1)}\ I_{2n+1}, & n=n' \end{cases},$$

where I_n denotes the identity matrix of order n. So, from (8.9), it follows that

$$(8.10)\qquad \frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH}) = \mathrm{tr}[M^{*-1}(\delta_{SH})\ D^*_{SH}],$$

where

$$(8.11)\qquad D^*_{SH} = \mathrm{diag}\left[1;\ \frac{1}{3},\frac{1}{3},\frac{1}{3};\ \frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5},\frac{1}{5};\ \ldots;\ 1/(2d+1),\ \ldots,\ 1/(2d+1)\right].$$

Thus, $\frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH})$ is a function only of the diagonal elements of $M^{*-1}(\delta_{SH})$.

Similarly, for F. S. regression of order d, we are searching for a δ^*_{FS} which minimizes

$$\frac{N}{\sigma^2} V_{\delta_{FS}}(w^*_{FS}) = \frac{1}{2\pi}\ \mathrm{tr}\left[M^{-1}(\delta_{FS})\int_0^{2\pi} \underset{\sim}{f}(\phi)\ \underset{\sim}{f}'(\phi)\ d\phi\right],$$

where

$$(8.12)\qquad M(\delta_{FS}) = \textstyle\sum_{i=1}^{p} w_i\ \underset{\sim}{f}(\phi_i)\ \underset{\sim}{f}'(\phi_i)$$

and where integration can be looked upon as being performed with respect to the uniform weighting function $w^*_{FS}(\phi) = \phi/2\pi$ either on the circumference of a circle or on the closed interval $[0,2\pi]$. By appealing to (2.2) and to the orthogonality properties of the "Fourier sequence" and by noting, from (1.6), that $\underset{\sim}{f}(\phi)\ \underset{\sim}{f}'(\phi)$ is a $(2d+1) \times (2d+1)$ block matrix with typical element $\underset{\sim}{f}_{(n)}(\phi)\ \underset{\sim}{f}'_{(n')}(\phi)$, $n = 0, 1, \ldots, d$, $n' = 0, 1, \ldots, d$, it follows that

$$(8.13) \qquad \frac{1}{2\pi}\int_0^{2\pi} \underset{\sim}{f}_{(n)}(\phi)\ \underset{\sim}{f}'_{(n')}(\phi)\ d\phi = \begin{cases} 0, & n \neq n' \\ 1, & n = n' = 0 \\ \frac{1}{2} I_2, & n = n' > 0 \end{cases},$$

and hence that

$$(8.14) \qquad \frac{N}{\sigma^2} V_{\delta_{FS}}(w^*_{FS}) = \mathrm{tr}[M^{-1}(\delta_{FS})\ D_{FS}],$$

where

$$(8.15) \qquad D_{FS} = \mathrm{diag}\left[1;\ \tfrac{1}{2},\tfrac{1}{2};\ \tfrac{1}{2},\tfrac{1}{2};\ \ldots;\ \tfrac{1}{2},\tfrac{1}{2}\right].$$

Note that (8.14) involves only the diagonal elements of $M^{-1}(\delta_{FS})$.

The implications of the following two remarks will be quite useful.

Remark 8.1. The set of $(2d+1)$ integers $\{0, 1, \ldots, 2d\}$ can be partitioned into $(d+1)$ disjoint sets $A_0, A_1, \ldots, A_d$, where $A_0 = \{0\}$ and $A_n = \{2n-1, 2n\}$ for $n = 1, 2, \ldots, d$.

Remark 8.2. The set of $(d+1)^2$ integers $\{0, 1, \ldots, d(d+2)\}$ can be partitioned into $(d+1)$ disjoint sets $B_0, B_1, \ldots, B_d$, where $B_0 = \{0\}$ and $B_n = \{n^2, n^2 + 2m - 1$ for $m = 1, 2, \ldots, n$, and $n^2 + 2m$ for $m = 1, 2, \ldots, n\}$, $n = 1, 2, \ldots, d$. Note that B_n has $(2n+1)$ elements, $n = 0, 1, \ldots, d$.

If we write $M^*(\delta_{SH}) = ((m^*_{i,j}(\delta_{SH})))_{i,j=0}^{d(d+2)}$, where $M^*(\delta_{SH})$ is given by (8.8), then the form of (8.4) and the partition discussed in Remark 8.2 allow us to make the following identifications:

$$(8.16) \qquad m^*_{n'^2,n^2}(\delta_{SH}) = \textstyle\sum_{i=1}^{p} w_i\ P_{n'}(\cos\theta_i)\ P_n(\cos\theta_i),$$

(8.17) $$m^*_{n'^2,n^2+2m-1}(\delta_{SH})$$

$$= C_{nm}\sum_{i=1}^{p} w_i\, P_{n'}(\cos\theta_i)\cos m\phi_i\, P_{nm}(\cos\theta_i),$$

(8.18) $$m^*_{n'^2,n^2+2m}(\delta_{SH})$$

$$= C_{nm}\sum_{i=1}^{p} w_i\, P_{n'}(\cos\theta_i)\sin m\phi_i\, P_{nm}(\cos\theta_i),$$

(8.19) $$m^*_{n^2+2m-1,n'^2+2m'-1}(\delta_{SH})$$

$$= C_{nm}\, C_{n'm'}\sum_{i=1}^{p} w_i \cos m\phi_i\, P_{nm}(\cos\theta_i)\cos m'\phi_i\, P_{n'm'}(\cos\theta_i),$$

(8.20) $$m^*_{n^2+2m-1,n'^2+2m'}(\delta_{SH})$$

$$= C_{nm}\, C_{n'm'}\sum_{i=1}^{p} w_i \cos m\phi_i\, P_{nm}(\cos\theta_i)\sin m'\phi_i\, P_{n'm'}(\cos\theta_i),$$

(8.21) $$m^*_{n^2+2m,n'^2+2m'}(\delta_{SH})$$

$$= C_{nm}\, C_{n'm'}\sum_{i=1}^{p} w_i \sin m\phi_i\, P_{nm}(\cos\theta_i)\sin m'\phi_i\, P_{n'm'}(\cos\theta_i),$$

where $C_{nm} = \left(\frac{2(n-m)!}{(n+m)!}\right)^{1/2}$.

In a similar manner, a labeling of the elements of $M(\delta_{FS}) = ((m_{i,j}(\delta_{FS})))_{i,j=0}^{2d}$ can be made by using the partitioning of Remark 8.1 and by noting the form of (1.6). In fact, it is easy to see that the following relations hold:

(8.22) $$m_{0,0}(\delta_{FS}) = 1,$$

(8.23) $$m_{0,2n-1}(\delta_{FS}) = \sum_{i=1}^{p} w_i \cos n\phi_i,$$

(8.24) $$m_{0,2n}(\delta_{FS}) = \sum_{i=1}^{p} w_i \sin n\phi_i,$$

(8.25) $$m_{2n-1,2n'-1}(\delta_{FS}) = \sum_{i=1}^{p} w_i \cos n\phi_i \cos n'\phi_i,$$

(8.26) $$m_{2n-1,2n'}(\delta_{FS}) = \sum_{i=1}^{p} w_i \cos n\phi_i \sin n'\phi_i,$$

(8.27) $$m_{2n,2n'}(\delta_{FS}) = \sum_{i=1}^{p} w_i \sin n\phi_i \sin n'\phi_i.$$

For notational convenience, we shall sometimes write $m^*_{i,j}$ for $m^*_{i,j}(\delta_{SH})$ and $m_{i,j}$ for $m_{i,j}(\delta_{FS})$.

Using (8.16), (8.19), (8.21) and (3.9), we can show after a little work that the following relations hold among the diagonal elements of $M^*(\delta_{SH})$:

(8.28) $$m^*_{0,0} = 1;\ m^*_{n^2,n^2} + \sum_{m=1}^{n} (m^*_{n^2+2m-1,n^2+2m-1} + m^*_{n^2+2m,n^2+2m}) = 1,$$

$$n = 1, 2, \ldots, d.$$

And, from (8.22), (8.25) and (8.27), it is obvious that

(8.29) $$m_{0,0} = 1;\ m_{2n-1,2n-1} + m_{2n,2n} = 1,\ n = 1, 2, \ldots, d.$$

Now, the form of (8.11) permits us to express (8.10) as

(8.30) $$\frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH}) = m^{*0,0} + \sum_{n=1}^{d} \frac{1}{(2n+1)} \left[m^{*n^2,n^2} + \sum_{m=1}^{n} (m^{*n^2+2m-1,n^2+2m-1} + m^{*n^2+2m,n^2+2m}) \right],$$

where $M^{*-1}(\delta_{SH}) = ((m^{*i,j}(\delta_{SH})))_{i,j=0}^{d(d+2)}$.

Similarly, if $M^{-1}(\delta_{FS}) = ((m^{i,j}(\delta_{FS})))_{i,j=0}^{2d}$, then, from (8.15), we can put (8.14) in the form

(8.31) $$\frac{N}{\sigma^2} V_{\delta_{FS}}(w^*_{FS}) = m^{0,0} + \frac{1}{2}\sum_{n=1}^{d} (m^{2n-1,2n-1} + m^{2n,2n}).$$

From the structures of the expressions (8.30) and (8.31), it follows that we can minimize $\frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH})$ and $\frac{N}{\sigma^2} V_{\delta_{FS}}(w^*_{FS})$ by properly specifying the values of the diagonal elements of $M^{*-1}(\delta_{SH})$ and $M^{-1}(\delta_{FS})$, respectively. However, the optimal choices for these elements cannot be made freely, due to the relations (8.28) and (8.29). To take these restrictions into account, we use Lagrange's technique and attempt to minimize by proper choices of $M^*(\delta_{SH})$ and $M(\delta_{FS})$ the quantities

$$(8.32)\quad f(\delta_{SH}) = m^{*0,0} + \sum_{n=1}^{d} \frac{1}{(2n+1)}\Big[m^{*n^2,n^2} + \sum_{m=1}^{n}(m^{*n^2+2m-1,n^2+2m-1} + m^{*n^2+2m,n^2+2m})\Big] + \lambda_0^*(m^*_{0,0}-1)$$
$$+ \sum_{n=1}^{d}\lambda_n^*\Big[m^*_{n^2,n^2} + \sum_{m=1}^{n}(m^*_{n^2+2m-1,n^2+2m-1} + m^*_{n^2+2m,n^2+2m}) - 1\Big],$$

and

$$(8.33)\quad f(\delta_{FS}) = m^{0,0} + \tfrac{1}{2}\sum_{n=1}^{d}(m^{2n-1,2n-1} + m^{2n,2n}) + \lambda_0(m_{0,0}-1)$$
$$+ \sum_{n=1}^{d} \lambda_n(m_{2n-1,2n-1} + m_{2n,2n} - 1),$$

respectively. The λ^*'s and λ's appearing in (8.32) and (8.33) are Lagrange multipliers.

Before we proceed with the minimization of $f(\delta_{SH})$ and $f(\delta_{FS})$, the following result will be needed.

Remark 8.3. Let $A = ((a_{ij}))_{i,j=1}^n$, $A^{-1} = ((a^{ij}))_{i,j=1}^n$, where A is taken to be symmetric and non-singular. From $A^{-1}A = I_n$, we have $(\partial A^{-1}/\partial a^{\alpha\beta})\, A + A^{-1}(\partial A/\partial a^{\alpha\beta}) = 0$, where $(\partial A/\partial a^{\alpha\beta}) = ((\partial a_{ij}/\partial a^{\alpha\beta}))_{i,j=1}^n$.

Thus, it follows that

$$\partial A/\partial a^{\alpha\beta} = -A(\partial A^{-1}/\partial a^{\alpha\beta})A = \begin{cases} -AE_{\alpha\alpha}A, & \alpha = \beta \\ -A(E_{\alpha\beta}+E_{\beta\alpha})A, & \alpha \neq \beta \end{cases},$$

where $E_{\alpha\beta}$ is the n×n matrix with a one in the (α,β)th position and zeros everywhere else. More specifically, if $\underset{\sim}{a}'_\alpha = (a_{\alpha 1}, a_{\alpha 2}, \ldots, a_{\alpha n})$ represents the αth row of A, then $\partial A/\partial a^{\alpha\alpha} = -\underset{\sim}{a}_\alpha\, \underset{\sim}{a}'_\alpha$, or $\partial a_{ij}/\partial a^{\alpha\alpha} = -a_{\alpha i}\, a_{\alpha j}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, n$.

Using (8.32) and appealing to Remark 8.3, we have:

(8.34) $$\partial f(\delta_{SH})/\partial m^{*n^2,n^2} = 1/(2n+1) - \lambda_0^* m^{*2}_{n^2,0} - \sum_{n'=1}^{d} \lambda_{n'}^* \Big[m^{*2}_{n^2,n'^2} + \sum_{m'=1}^{n'} (m^{*2}_{n^2,n'^2+2m'-1} + m^{*2}_{n^2,n'^2+2m'}) \Big],$$

$n = 0, 1, \ldots, d;$

(8.35) $$\partial f(\delta_{SH})/\partial m^{*n^2+2m-1,n^2+2m-1} = 1/2n+1) - \lambda_0^* m^{*2}_{n^2+2m-1,0} - \sum_{n'=1}^{d} \lambda_{n'}^* \Big[m^{*2}_{n^2+2m-1,n'^2} + \sum_{m'=1}^{n'} (m^{*2}_{n^2+2m-1,n'^2+2m'-1} + m^{*2}_{n^2+2m-1,n'^2+2m'}) \Big],$$ $m = 1, 2, \ldots, n$, $n = 1, 2, \ldots, d;$

and,

(8.36) $$\partial f(\delta_{SH})/\partial m^{*n^2+2m,n^2+2m} = 1/(2n+1) - \lambda_0^* m^{*2}_{n^2+2m,0} - \sum_{n'=1}^{d} \lambda_{n'}^* \Big[m^{*2}_{n^2+2m,n'^2} + \sum_{m'=1}^{n'} (m^{*2}_{n^2+2m,n'^2+2m'-1} + m^{*2}_{n^2+2m,n'^2+2m'}) \Big],$$ $m = 1, 2, \ldots, n$, $n = 1, 2, \ldots, d.$

Equating the $(d+1)^2$ partial derivatives (8.34) - (8.36) to zero and solving the resulting set of non-linear equations in conjunction with the (d+1) relations given by (8.28), we obtain the following solution which characterizes the WV design δ^*_{SH} in terms of $M^*(\delta^*_{SH})$:

(8.37) $m^*_{0,0} = 1$; $m^*_{n^2,n^2} = m^*_{n^2+2m-1,n^2+2m-1} = m^*_{n^2+2m,n^2+2m} = 1/(2n+1)$,

$m = 1, 2, \ldots, n$, $n = 1, 2, \ldots, d$; $m^*_{i,j} = 0$ if $i \neq j$;

and, $\lambda^*_n = (2n+1)$, $n = 0, 1, \ldots, d$.

Similarly, it follows from (8.33) and Remark 8.3 that

(8.38) $\partial f(\delta_{FS})/\partial m^{0,0} = 1 - \lambda_0 m^2_{0,0} - \sum^d_{n=1} \lambda_n(m^2_{0,2n-1} + m^2_{0,2n})$;

(8.39) $\partial f(\delta_{FS})/\partial m^{2n-1,2n-1} = \frac{1}{2} - \lambda_0 m^2_{2n-1,0} - \sum^d_{n'=1} \lambda_{n'}(m^2_{2n-1,2n'-1}$

$+ m^2_{2n-1,2n'})$, $n = 1, 2, \ldots, d$;

and,

(8.40) $\partial f(\delta_{FS})/\partial m^{2n,2n} = \frac{1}{2} - \lambda_0 m^2_{2n,0} - \sum^d_{n'=1} \lambda_{n'}(m^2_{2n,2n'-1} + m^2_{2n,2n'})$,

$n = 1, 2, \ldots, d$.

As before, we set the (2d+1) partial derivatives (8.38) - (8.40) equal to zero and solve along with (8.29) to obtain the specification of δ^*_{FS} in terms of $M(\delta^*_{FS})$; namely,

(8.41) $m_{0,0} = 1$; $m_{2n-1,2n-1} = m_{2n,2n} = 1/2$, $n = 1, 2, \ldots, d$; $m_{i,j} = 0$

if $i \neq j$; and, $\lambda_0 = 1$, $\lambda_n = 2$, $n = 1, 2, \ldots, d$.

To conclude, as we have done, that the designs δ^*_{SH} and δ^*_{FS}, characterized in terms of their information matrices $M^*(\delta^*_{SH})$ and $M(\delta^*_{FS})$

by the solutions (8.37) and (8.41), respectively, are indeed WV designs, it is necessary to show that the expressions (8.30) and (8.31) are minimized, subject to the restrictions (8.28) and (8.29), at δ^*_{SH} and δ^*_{FS}.

To do this, we proceed as follows. Let the $(d+1)^2 \times (d+1)^2$ matrix $F^*_{SH} = ((f^*_{ij}))^{d(d+2)}_{i,j=0}$ have as its typical element the quantity $f^*_{ij} = (\partial^2 f(\delta_{SH})/\partial m^{*ii}\ \partial m^{*jj})_{(8.37)}$ (i.e., the second partial derivative of $f(\delta_{SH})$ evaluated at the solution (8.37)), and let the $(d+1)^2 \times (d+1)$ matrix $G^*_{SH} = ((g^*_{in}))^{d(d+2),d}_{i=0,n=0}$ be defined such that $g^*_{in} = (\partial^2 f(\delta_{SH})/\partial m^{*ii}\ \partial \lambda^*_n)_{(8.37)}$. Then, a necessary and sufficient condition for $\frac{N}{\sigma^2} V_{\delta_{SH}}(w^*_{SH})$ to be minimized at δ^*_{SH} subject to the conditions (8.28) is that the roots of the d(d+1)th degree polynomial in λ obtained by expanding $\begin{vmatrix} F^*_{SH}-\lambda I & G^*_{SH} \\ G^{*\prime}_{SH} & 0 \end{vmatrix} = 0$ are <u>all</u> positive. (See Hancock [16], Chapter VI). Similarly, $\frac{N}{\sigma^2} V_{\delta_{FS}}(w^*_{FS})$ achieves its minimum at δ^*_{FS} subject to the conditions (8.29) if the roots of the dth degree polynomial in λ obtained by expanding $\begin{vmatrix} F_{FS}-\lambda I & G_{FS} \\ G'_{FS} & 0 \end{vmatrix} = 0$ are <u>all</u> positive, where the $(2d+1) \times (2d+1)$ matrix $F_{FS} = ((f_{ij}))^{2d}_{i,j=0}$ has as its typical element $f_{ij} = (\partial^2 f(\delta_{FS})/\partial m^{ii}\ \partial m^{jj})_{(8.41)}$ and where the $(2d+1) \times (d+1)$ matrix $G_{FS} = ((g_{in}))^{2d,d}_{i=0,n=0}$ has as its typical element $g_{in} = (\partial^2 f(\delta_{FS})/\partial m^{ii}\ \partial \lambda_n)_{(8.41)}$.

Now, the determination of the elements of the matrices F^*_{SH}, F_{FS}, G^*_{SH} and G_{FS} is tedious but straightforward. For example, from (8.34) and Remark 8.3, it follows directly that $\partial^2 f(\delta_{SH})/\partial m^*{}_{n^2,n^2}\, \partial m^*{}_{n'^2,n'^2} =$

$$2\lambda_0^* m^*_{n^2,0}\, m^*_{n'^2,n^2}\, m^*_{n'^2,0} + 2\sum_{n''=1}^{d} \lambda^*_{n''}\Big[m^*_{n^2,n''^2}\, m^*_{n'^2,n^2}\, m^*_{n'^2,n''^2}$$

$$+ \sum_{m''=1}^{n''}(m^*_{n^2,n''^2+2m''-1}\, m^*_{n'^2,n^2}\, m^*_{n'^2,n''^2+2m''-1}$$

$$+ m^*_{n^2,n''^2+2m''}\, m^*_{n'^2,n^2}\, m^*_{n'^2,n''^2+2m''})\Big],$$

and so, using (8.37), we have $f^*_{n^2,n'^2} = \begin{cases} 0 \text{ if } n \neq n' \\ 2/(2n+1)^2 \text{ if } n=n' \end{cases}$ for $0 \leq n, n' \leq d$. It is just as straightforward to show that, in general, $f^*_{0,0} = 2$, $f^*_{n^2,n^2} = f^*_{n^2+2m-1,n^2+2m-1} = f^*_{n^2+2m,n^2+2m} = 2/(2n+1)^2$, $m = 1, 2, \ldots, n$, $n = 1, 2, \ldots, d$, and $f^*_{ij} = 0$ if $i \neq j$. In other words, $F^*_{SH} = 2D^{*2}_{SH}$, where D^*_{SH} is given by (8.11). Also, from (8.34) - (8.36), it is readily verified that $g^*_{0,0} = -1$ and $g^*_{i,0} = 0$ for $i = 1, 2, \ldots, d(d+2)$, and that, for each fixed n, $n = 1, 2, \ldots, d$, $g^*_{n^2,n} = g^*_{n^2+2m-1,n} = g^*_{n^2+2m,n} = -1/(2n+1)^2$ for $m = 1, 2, \ldots, n$ and $g^*_{in} = 0$ for all other i. Thus, it follows that $2/(2n+1)^2$ is a root of multiplicity $2n$ of

$$\begin{vmatrix} F^*_{SH}-\lambda I & G^*_{SH} \\ G^{*\prime}_{SH} & 0 \end{vmatrix} = 0$$

for $n = 1, 2, \ldots, d$; since these $d(d+1)$ roots are all positive, we can conclude that δ^*_{SH} is a WV design for S. H. regression of order d. Similarly, using (8.38) - (8.40), we can show that $F_{FS} = 2D^2_{FS}$, where D_{FS} is given by (8.15); also, it is clear that

$g_{0,0} = -1$ and $g_{i,0} = 0$ for $i = 1, 2, \ldots, 2d$, and that, for each fixed n, $n = 1, 2, \ldots, d$, $g_{2n-1,n} = g_{2n,n} = -1/4$ and $g_{in} = 0$ for all other i. Since it follows from these results that 1/2 is a root of multiplicity d of $\begin{vmatrix} F_{FS}-\lambda I & G_{FS} \\ G'_{FS} & 0 \end{vmatrix} = 0$, δ^*_{FS} is a WV design for F. S. regression of order d.

Using (8.37), we can then say that a WV design δ^*_{SH} for S. H. regression of order d can be characterized as one whose information matrix has the form $M^*(\delta^*_{SH}) = D^*_{SH}$, where $M^*(\delta_{SH})$ is given by (8.8) and D^*_{SH} by (8.11); also, from (8.41), it follows that a WV design δ^*_{FS} for F. S. regression of order d has the information matrix $M(\delta^*_{FS}) = D_{FS}$, where $M(\delta_{FS})$ is given by (8.12) and D_{FS} by (8.15).

We can now appeal directly to Lemma 5.2 to show that δ^*_{SH} and δ^*_{FS} are also MV designs, since, from (8.6),

$$(8.42) \quad \underset{\sim}{f}^{*\prime}(\theta,\phi)\, D^{*-1}_{SH}\, \underset{\sim}{f}^{*}(\theta,\phi) = \textstyle\sum_{n=0}^{d}(2n+1)\, \underset{\sim}{f}'_{[n]}(\theta,\phi)\, \underset{\sim}{f}_{[n]}(\theta,\phi)$$

$$= \textstyle\sum_{n=0}^{d}(2n+1) = (d+1)^2,$$

and, from (1.8),

$$(8.43) \quad \underset{\sim}{f}'(\phi)\, D^{-1}_{FS}\, \underset{\sim}{f}(\phi) = \underset{\sim}{f}'_{(0)}(\phi)\, \underset{\sim}{f}_{(0)}(\phi) + 2\textstyle\sum_{n=1}^{d} \underset{\sim}{f}'_{(n)}(\phi)\, \underset{\sim}{f}_{(n)}(\phi)$$

$$= 1 + 2\textstyle\sum_{n=1}^{d}(1) = (2d+1).$$

Also, we must have

$$(8.44) \quad \frac{N}{\sigma^2} V_{\delta^*_{SH}}(w^*_{SH}) = (d+1)^2 \text{ and } \frac{N}{\sigma^2} V_{\delta^*_{FS}}(w^*_{FS}) = (2d+1);$$

(8.44) is obvious from the form of expressions (8.10) and (8.14).

Thus, we have been able to characterize, in terms of their corresponding information matrices, the optimal approximate designs δ^*_{SH} and δ^*_{FS} for S. H. and F. S. regression of order d which simultaneously minimize both a weighted average variance and maximum variance over certain specified regions.

The general discussion given in Section 6 will now be used to show that δ^*_{SH} and δ^*_{FS} are also WB designs. In particular, suppose that the S. H. model to be fitted by least squares is of order d, while the true response is $\eta_{d'}(\theta,\phi;\underset{\sim}{\beta}_{SH})$, where $d' > d \geq 1$. Then, in this setting, we know from the results of Section 6 that a WB design can be characterized as one which makes the $(d+1)^2 \times (d+1)^2$ matrix $M^*_{11}(\delta_{SH})$ (or, equivalently, $M^*(\delta_{SH})$) equal to Λ^*_{11} (again, no problem will result from working with $M^*(\delta_{SH})$ instead of $M(\delta_{SH})$) and which makes the $(d+1)^2 \times (d'-d)(d'+d+2)$ matrix $M^*_{12}(\delta_{SH}) = \sum_{i=1}^{p} w_i \, \underset{\sim}{f}^*(\theta_i,\phi_i) \, \underset{\sim}{f}_B^{*\prime}(\theta_i,\phi_i)$ take the form Λ^*_{12}, where $M^*(\delta_{SH})$ is given by (8.8),

$$\underset{\sim}{f}_B^{*\prime}(\theta,\phi) = \left(\underset{\sim}{f}'_{[d+1]}(\theta,\phi), \ \ldots, \ \underset{\sim}{f}'_{[d']}(\theta,\phi)\right),$$

$$\Lambda^*_{11} = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \underset{\sim}{f}^*(\theta,\phi) \, \underset{\sim}{f}^{*\prime}(\theta,\phi) \sin\theta \, d\theta d\phi,$$

$$\Lambda^*_{12} = \frac{1}{4\pi} \int_0^{2\pi} \int_0^{\pi} \underset{\sim}{f}^*(\theta,\phi) \, \underset{\sim}{f}_B^{*\prime}(\theta,\phi) \sin\theta \, d\theta d\phi,$$

and where the integrals are to be given the same interpretations here as they were for the WV design problem discussed earlier in this section. Now, it is easy to see from (8.9) that $\Lambda^*_{11} = D^*_{SH}$ and $\Lambda^*_{12} = 0$, where D^*_{SH} is given by (8.11). Hence, the WB design for S. H. regression, which is only appropriate when the bias error is due solely to the fitting (by least squares) of an S. H. model of too small a degree, is simply the WV and MV design δ^*_{SH}, characterized earlier in this section, which has its associated information matrix $M^*(\delta^*_{SH}) = D^*_{SH}$ of "large enough size" to insure that the orthogonality conditions implied by the matrix equation $M^*_{12}(\delta^*_{SH}) = 0$ are satisfied. This last notion will be made more precise in the next section.

In a completely analogous manner, it follows that a WB design δ^*_{FS} for F. S. regression has an information matrix of the general form D_{FS}, since we are to take $M_{11}(\delta_{FS}) \equiv M(\delta_{FS})$ equal to

$\frac{1}{2\pi}\int_0^{2\pi} \underset{\sim}{f}(\phi)\ \underset{\sim}{f}'(\phi)\ d\phi = D_{FS}$ and $M_{12}(\delta_{FS}) = \sum_{i=1}^{p} w_i\ \underset{\sim}{f}(\phi_i)\ \underset{\sim}{f}'_B(\phi_i)$

equal to $\frac{1}{2\pi}\int_0^{2\pi} \underset{\sim}{f}(\phi)\ \underset{\sim}{f}'_B(\phi)\ d\phi = 0$, where $\underset{\sim}{f}'_B(\phi) = \Big(\underset{\sim}{f}'_{(d+1)}(\phi),$

$\ldots, \underset{\sim}{f}'_{(d')}(\phi)\Big)$. Again, D_{FS} must be chosen so that the orthogonality requirements given by $M_{12}(\delta^*_{FS}) = 0$ are fulfilled. It is important to mention that the WB designs δ^*_{SH} and δ^*_{FS} also possess another desirable property which was discussed in Section 6; namely, the alias matrices $A^*(\delta^*_{SH}) = M^{*-1}(\delta^*_{SH})\ M^*_{12}(\delta^*_{SH})$ and $A(\delta^*_{FS})$ $= M^{-1}(\delta^*_{FS})\ M_{12}(\delta^*_{FS})$ are null matrices.

Since δ^*_{SH} and δ^*_{FS} are also WMS designs in the above framework, it follows directly from Theorem 7.1 that they are each simultaneously V-, B-, and MS- admissible. And, by appealing to Theorem 7.1, p. 808, of [19] and to the relations (8.42) and (8.43), we can show fairly directly that δ^*_{SH} and δ^*_{FS} are also admissible in the sense of Definition 7.1.

Now, the optimal properties of δ^*_{SH} and δ^*_{FS} that we have discussed so far are relevant mainly in a response surface setting and that is why they are concerned with estimation of the entire regression function rather than of specific parameters. In what now follows, we will show that the information matrices D^*_{SH} and D_{FS} also possess several important properties which are intimately related to the optimal estimation of individual regression coefficients in the S. H. and F. S. models and of meaningful functions of the coefficients.

First of all, since δ^*_{SH} and δ^*_{FS} are MV designs (or "G-optimal" designs in approximate theory terminology), the Equivalence Theorem (see [23]) mentioned in Section 5 allows us to conclude that $|M^*(\delta_{SH})|$ and $|M(\delta_{FS})|$ are maximized at δ^*_{SH} and at δ^*_{FS}, respectively; in other words, δ^*_{FS} and δ^*_{SH} are both "D-optimal".

Before considering some other design criteria, let us digress for just a moment to make the following remarks. In all the work that we have done in this section up to now, the optimal approximate design δ^*_{SH} has been characterized using $M^*(\delta_{SH})$ instead of $M(\delta_{SH})$ $= \sum_{i=1}^{p} w_i \, \underset{\sim}{f}(\theta_i,\phi_i) \, \underset{\sim}{f}'(\theta_i,\phi_i)$, where

(8.45) $$M^*(\delta_{SH}) = G_{SH}^{1/2} \, M(\delta_{SH}) \, G_{SH}^{1/2},$$

G_{SH} being the $(d+1)^2 \times (d+1)^2$ diagonal matrix

(8.46) $$G_{SH} = \text{diag}(\underset{\sim}{g}'_0, \underset{\sim}{g}'_1, \ldots, \underset{\sim}{g}'_d),$$

where

(8.47) $$\underset{\sim}{g}_0 = 1; \; \underset{\sim}{g}'_n = \left[1, \frac{2}{n(n+1)}, \frac{2}{n(n+1)}, \ldots, \frac{2(n-m)!}{(n+m)!}, \frac{2(n-m)!}{(n+m)!}, \ldots, \frac{2}{(2n)!}, \frac{2}{(2n)!}\right], \; n = 1, 2, \ldots, d.$$

The expression (8.45) giving the relationship between $M^*(\delta_{SH})$ and $M(\delta_{SH})$ follows directly by noting from (8.1) that

(8.48) $$\underset{\sim}{f}^*(\theta,\phi) = G_{SH}^{1/2}\, \underset{\sim}{f}(\theta,\phi).$$

Similarly, from (8.2), we can write

(8.49) $$\underset{\sim}{\beta}^*_{SH} = G_{SH}^{-1/2}\, \underset{\sim}{\beta}_{SH}.$$

Now, one can easily verify that the design criteria for S. H. regression so far considered are invariant under the pair of transformations (8.48) and (8.49) in the sense that, for any of these criteria, the optimal design δ^*_{SH} characterized by $M^*(\delta^*_{SH})$ is exactly the same as the optimal design δ^o_{SH} characterized by $M(\delta^o_{SH})$. The reason we bring this point up at this particular time is that the design criteria now about to be discussed do not possess this desirable invariance property and hence, for these criteria, the designs δ^*_{SH} and δ^o_{SH} will, in general, be different. In what follows, we shall continue as before to work with $M^*(\delta_{SH})$.

Techniques completely analogous to those used earlier in finding WV designs for S. H. and F. S. regression can be applied to show that δ^*_{SH} and δ^*_{FS} are both "A-optimal" designs; in other words, tr $M^{*-1}(\delta_{SH})$ and tr $M^{-1}(\delta_{FS})$ are minimized at δ^*_{SH} and at δ^*_{FS}, respectively. In brief, the verification of the A- optimality of δ^*_{SH} and δ^*_{FS} follows directly by using the differentiation formulae of Remark 8.3 to minimize the quantities tr $M^{*-1}(\delta_{SH}) = m^{*0,0} + \sum_{n=1}^{d}\Big[m^{*n^2,n^2} + \sum_{m=1}^{n}(m^{*n^2+2m-1,n^2+2m-1} + m^{*n^2+2m,n^2+2m})\Big]$ and tr $M^{-1}(\delta_{FS}) = m^{0,0} + \sum_{n=1}^{d}(m^{2n-1,2n-1} + m^{2n,2n})$ subject to the restrictions (8.28) and (8.29), respectively, which are introduced via Lagrangian multipliers as in (8.32) and (8.33).

Now, let $M(\delta)$ be the information matrix corresponding to some $\delta \varepsilon \Delta$, where $M(\delta)$ has the typical element $m_{ij}(\delta) \equiv m_{ij}$; similarly, let $m^{ij}(\delta) \equiv m^{ij}$ be the (i,j)th entry in $M^{-1}(\delta)$. In this framework, we give the following definition (see Elfving, [12]).

Definition 8.1. A design $\delta^* \varepsilon \Delta$ is said to be m.s.p. (minimax with respect to single parameters) when it minimizes $\mathrm{MAX}_i m^{ii}$ on Δ; it is said to be m.st.f. (minimax with respect to standard parametric forms) when it minimizes $\underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\mathrm{MAX}} \underset{\sim}{c}' M^{-1}(\delta) \underset{\sim}{c}$ on Δ.

When all the parameters in the fitted model are to be included in the minimax approach defined above, we shall refer to this as "total minimaxity"; when only a subset of these parameters is of interest, we will be concerned with "partial minimaxity".

In the above setting, let h be a subscript for which $m_{hh} = \mathrm{MIN}_i m_{ii}$. Then, since $m^{hh} \geq (m_{hh})^{-1}$, we have

$$(8.50) \qquad \mathrm{MAX}_i m^{ii} \geq m^{hh} \geq (m_{hh})^{-1} \geq k/\mathrm{tr}\, M(\delta),$$

where $M(\delta)$ is taken to be $(k \times k)$. Clearly, a sufficient condition for equality is that $M(\delta) = \lambda I_k$ for some $\lambda > 0$.

If $\lambda(M)$ represents a characteristic root of $M(\delta)$, then it follows that

$$(8.51) \quad \underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\mathrm{MAX}} \underset{\sim}{c}' M^{-1}(\delta) \underset{\sim}{c} = \lambda_{max}(M^{-1}) = 1/\lambda_{min}(M) \geq k/\mathrm{tr}\, M(\delta);$$

again, a sufficient condition for equality is that $M(\delta)$ be of the form λI_k for some constant $\lambda > 0$.

In those situations when $\mathrm{tr}\, M(\delta)$ is either constant on Δ or has an obvious upper bound, it is clear that (8.50) and (8.51) will then provide lower bounds for $\text{MAX}_i m^{ii}$ and $\underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\text{MAX}} \underset{\sim}{c}' M^{-1}(\delta) \underset{\sim}{c}$, respectively.

Since $\text{tr } M^*(\delta_{SH}) = \text{tr } M(\delta_{FS}) = (d+1)$, it follows that the quantities $\text{MAX}_i m^{ii}$ and $\underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\text{MAX}} \underset{\sim}{c}' M^{-1}(\delta) \underset{\sim}{c}$ have a lower bound of $(d+1)$ for $M^*(\delta_{SH})$ and a lower bound of $[1 + \frac{d}{(d+1)}]$ for $M(\delta_{FS})$. Except in the uninteresting case $d = 0$, these lower bounds are not attained by δ^*_{SH} and δ^*_{FS}, since for $d \geq 1$, $\text{MAX}_i m^{*ii}(\delta^*_{SH}) = \underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\text{MAX}} \underset{\sim}{c}' M^{*-1}(\delta^*_{SH}) \underset{\sim}{c} = (2d+1)$ and $\text{MAX}_i m^{ii}(\delta^*_{FS}) = \underset{\underset{\sim}{c},\underset{\sim}{c}'\underset{\sim}{c}=1}{\text{MAX}} \underset{\sim}{c}' M^{-1}(\delta^*_{FS}) \underset{\sim}{c} = 2$. In fact, the restrictions (8.28) and (8.29) prohibit us from choosing $M^*(\delta_{SH})$ and $M(\delta_{FS})$ to be multiples of identity matrices. Thus, the designs δ^*_{SH} and δ^*_{FS} are neither m.s.p. nor m.st.f. in the context of "total minimaxity" with respect to $\underset{\sim}{\beta}^*_{SH}$ and $\underset{\sim}{\beta}_{FS}$.

In the framework of "partial minimaxity", one can develop inequalities completely analogous to (8.50) and (8.51) by introducing the notion of a "partial information matrix". By appealing directly to Theorem 3.4 of Elfving [12], we may conclude that, for each fixed n, δ^*_{SH} is both m.s.p. and m.st.f. with respect to the $(2n+1)$ elements of $\underset{\sim}{\beta}^{[n]}_{SH}$ given by (8.2) and that δ^*_{FS} is both m.s.p. and m.st.f. with respect to the elements of $\underset{\sim}{\beta}^{(n)}_{FS}$ given in (1.4), $n = 1, 2, \ldots, d$. The partial information matrices corresponding to $\underset{\sim}{\beta}^{[n]}_{SH}$ and $\underset{\sim}{\beta}^{(n)}_{FS}$ are $\frac{1}{(2n+1)} I_{2n+1}$ and $\frac{1}{2} I_2$, respectively.

We may summarize the more important results of this section in the following theorem.

<u>THEOREM 8.1</u>. The approximate designs δ^*_{SH} and δ^*_{FS} characterized by

the information matrices $M^*(\delta^*_{SH}) = D^*_{SH}$ and $M(\delta^*_{FS}) = D_{FS}$, respectively, possess the following optimal properties:

(i) they are orthogonal designs (i.e., D^*_{SH} and D_{FS} are diagonal);

(ii) they are WV, WB, and WMS designs with respect to the least favorable weighting functions $w^*_{SH}(\theta,\phi) = (1-\cos\theta)\phi/4\pi$, $0 \leq \theta \leq \pi$, $0 \leq \phi \leq 2\pi$, and $w^*_{FS}(\phi) = \phi/2\pi$, $0 \leq \phi \leq 2\pi$ (see Remark 5.1);

(iii) they are V-, B-, and MS- admissible, and they are also admissible in the sense of Definition 7.1:

(iv) they are MV and D- optimal designs;

(v) they are A- optimal designs; and,

(vi) they are m.s.p. and m.st.f. with respect to the elements of $\underset{\sim}{\beta}^{[n]}_{SH}$ and $\underset{\sim}{\beta}^{(n)}_{FS}$, $n = 1, 2, \ldots, d$.

Conjecture: This author feels that δ^*_{SH} and δ^*_{FS} are also MB and MMS designs but has not been able to justify this belief.

Let us emphasize again that the design δ^*_{SH} characterized by $M^*(\delta^*_{SH}) = D^*_{SH}$ can be claimed to be identically the same as the design δ^o_{SH} characterized by $M(\delta^o_{SH}) = D_{SH} = G^{-1/2}_{SH}\ D^*_{SH}\ G^{-1/2}_{SH}$, where D^*_{SH} is given by (8.11) and where G_{SH} is given by the relations (8.46) and (8.47), only when referring to parts (i) - (iv) (and not to parts (v) and (vi)) of Theorem 8.1.

One final word is in order concerning the minimaxity criteria of Definition 8.1. Suppose that the model under study is of the form (4.1) and that there exists a design $\delta^* \in \Delta$ for which $M(\delta^*)$ is a diagonal matrix D, where $M(\delta) = \sum_{i=1}^{p} w_i\ \underset{\sim}{f}(\underset{\sim}{x}_i)\ \underset{\sim}{f}'(\underset{\sim}{x}_i)$ for any $\delta \in \Delta$. Then, if we write $\underset{\sim}{f}'(\underset{\sim}{x})\underset{\sim}{\beta} = (D^{-1/2}\underset{\sim}{f}(\underset{\sim}{x}))'(D^{1/2}\underset{\sim}{\beta}) = \underset{\sim}{f}^{*\prime}(\underset{\sim}{x})\underset{\sim}{\beta}^*$, it follows that δ^* is both m.s.p. and m.st.f. with respect to the elements of the parameter vector $\underset{\sim}{\beta}^* = D^{1/2}\underset{\sim}{\beta}$ since $M^*(\delta^*) = D^{-1/2}M(\delta^*)D^{-1/2} = I$, where

$M^*(\delta) = \sum_{i=1}^{p} w_i \underset{\sim}{f}^*(\underset{\sim}{x}_i) \underset{\sim}{f}^{*\prime}(\underset{\sim}{x}_i)$ for any $\delta \in \Delta$. This result implies, using previous notation, that δ^*_{SH} is m.s.p. and m.st.f. with respect to $D^{*1/2}_{SH} \underset{\sim}{\beta}^*_{SH} = D^{*1/2}_{SH} G^{-1/2}_{SH} \underset{\sim}{\beta}_{SH}$ and that δ^*_{FS} is m.s.p. and m.st.f. with respect to $D^{1/2}_{FS} \underset{\sim}{\beta}_{FS}$. From these comments, it is clear that one need only search for a diagonal information matrix to conclude that the associated design is both m.s.p. and m.st.f. with respect to <u>some</u> set of parameters.

9. The Characterization of δ^*_{SH} and δ^*_{FS} as Rotatable Arrangements

Consider the estimated response $\hat{y}_d(\underset{\sim}{x};\underset{\sim}{b}_p) = \underset{\sim}{f}'_d(\underset{\sim}{x}) \underset{\sim}{b}_p =$ $b_0 + b_1x_1 + b_2x_2 + \ldots + b_kx_k + b_{11}x_1^2 + \ldots + b_{kk}x_k^2 + b_{12}x_1x_2 +$ $\ldots + b_{k-1,k}x_{k-1}x_k + b_{111}x_1^3 + \ldots$, which is a polynomial of degree d in the k variables $x_1, x_2, \ldots, x_k$ obtained by the method of least squares using a design δ_p consisting of the N points $\underset{\sim}{x}'_i = (x_{1i}, x_{2i}, \ldots, x_{ki})$, $i = 1, 2, \ldots, N$.

If we write $M(\alpha_1, \alpha_2, \ldots, \alpha_k) = \frac{1}{N} \sum_{i=1}^{N} x_{1i}^{\alpha_1} x_{2i}^{\alpha_2} \ldots x_{ki}^{\alpha_k}$, then the set of design points $\{\underset{\sim}{x}_i\}_{i=1}^{N}$ is said to form a "rotatable arrangement" (see [5]) if it satisfies the moment conditions

$$(9.1)\quad M(\alpha_1, \alpha_2, \ldots, \alpha_k) = \begin{cases} 0 \text{ if one or more } \alpha_i \text{ are odd} \\[2ex] \dfrac{\lambda_\alpha \prod_{i=1}^{k} \alpha_i!}{2^{\alpha/2} \prod_{i=1}^{k} (\alpha_i/2)!} \text{ if all } \alpha_i \text{ are even} \end{cases},$$

where $0 \le \alpha_i \le 2d$ for every i and where $0 \le \alpha = \sum_{i=1}^{k} \alpha_i \le 2d$, α being the order of the moment; λ_α is a positive constant depending only on α. These moment conditions are equivalent to the condition that $\frac{N}{\sigma^2} \operatorname{Var} \hat{y}_d(\underset{\sim}{x};\underset{\sim}{b}_p) = \underset{\sim}{f}'_d(\underset{\sim}{x})\, M^{-1}(\delta_p)\, \underset{\sim}{f}_d(\underset{\sim}{x})$ be a function only of $\underset{\sim}{x}'\underset{\sim}{x}$, where

$M(\delta_P) = \frac{1}{N} \sum_{i=1}^{N} \underset{\sim}{f}_d(\underset{\sim}{x}_i)\ \underset{\sim}{f}_d'(\underset{\sim}{x}_i)$; however, they are not, in general, sufficient to make $M(\delta_P)$ non-singular, and, in this case, all the $\binom{k+d}{d}$ polynomial coefficients will not be individually estimable. In particular, when $d > 1$, this singular case will arise when all the design points are equidistant from the origin of the factor levels.

Using the notions just developed and the results of the previous section, we prove the following theorem.

THEOREM 9.1. The optimal approximate design δ_{SH}^* characterized by the $(d+1)^2 \times (d+1)^2$ information matrix D_{SH}^* is exactly that set of points $\{\theta_i, \phi_i\}_{i=1}^{N}$ with equal weights $w_i = 1/N$ for every i which satisfies the moment conditions (9.1) of a rotatable arrangement of order d when transformed to the set of points $\{\underset{\sim}{x}_{SH}^{(i)}\}_{i=1}^{N}$, where $\underset{\sim}{x}_{SH}$ is given by (3.4); similarly, the optimal approximate design δ_{FS}^* characterized in terms of the $(2d+1) \times (2d+1)$ information matrix D_{FS} is exactly that set of points $\{\phi_i\}_{i=1}^{N}$ with equal weights $w_i = 1/N$ for every i which satisfies the moment conditions (9.1) of a rotatable arrangement of order d when transformed to the set of points $\{\underset{\sim}{x}_{FS}^{(i)}\}_{i=1}^{N}$, where $\underset{\sim}{x}_{FS}$ is given by (3.3). In particular, then, it follows that designs δ_{SH}^* and δ_{FS}^* which are optimal for S. H. and F. S. regression of order d, respectively, are also optimal for all lower order regressions.

Proof. Let $\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})$ denote the estimated response at the point (θ,ϕ) obtained by the least squares fitting of an S. H. model of order d; further, let $\hat{y}_d(\underset{\sim}{x}_{SH};\underset{\sim}{b}_P)$ be a dth degree polynomial representation of the response $\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})$ at the point $\underset{\sim}{x}_{SH}$ corresponding to (θ,ϕ), where such a representation is described in Lemma 3.2 and Remark 3.1. It is obvious that the variance of $\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})$ at (θ,ϕ) must be exactly the same as the variance of $\hat{y}_d(\underset{\sim}{x}_{SH};\underset{\sim}{b}_P)$ at $\underset{\sim}{x}_{SH}$. Assuming that δ^*_{SH} has been used to fit the S. H. model, it follows from (8.42) that Var $\hat{y}_d(\underset{\sim}{x}_{SH};\underset{\sim}{b}_P)$ $= (d+1)^2\sigma^2/N$, which is independent of $\underset{\sim}{x}_{SH}$. Since $\underset{\sim}{x}'_{SH}\underset{\sim}{x}_{SH} = 1$, the design points $\{(\theta_i,\phi_i)\}_{i=1}^N$ must, by definition, form a rotatable arrangement of order d when expressed as the set of points $\{\underset{\sim}{x}_{SH}^{(i)}\}_{i=1}^N$. A completely analogous argument can be used to characterize δ^*_{FS} in a similar manner. Since a rotatable arrangement of order d is also a rotatable arrangement of any order $d' < d$, the last statement in the theorem is true. This completes the proof.

Now, there are some interesting comments to be made concerning the implications of the above theorem. First of all, the moments of a rotatable design of order d given by (9.1) are actually the moments up to order 2d of a general spherical distribution $p(\underset{\sim}{x}) = cf(\underset{\sim}{x}'\underset{\sim}{x})$, i.e., a distribution which is a function only of $\underset{\sim}{x}'\underset{\sim}{x} = \sum_{i=1}^k x_i^2$ and hence is constant (or "uniform") on hyperspheres centered at the origin of the x_i's. In this light, rotatable designs can be and are usually formed by combining one or more "rotatable arrangements" of different radii, a rotatable arrangement consisting of a set of points uniformly spaced on the surface of a hypersphere. An analytical argument advocating the use of an aesthetically appealing "rotatable arrangement" of points as a design in its own right has not been forthcoming until the present work.

Using the results concerning rotatable arrangements and designs that are already available in the literature, we will be able to specify optimal exact designs (δ^*_{FS}) for F. S. regression of any order d and to construct optimal exact designs (δ^*_{SH}) for S. H. regression of orders one and two. Essentially, the reason for the limited range of d in the latter case above is due to the fact that there are only five regular figures in three dimensions, none of which has vertices satisfying the moment conditions (9.1) for any $d \geq 3$. In fact, to paraphrase Kiefer [21], it is not generally possible to construct a rotatable arrangement of points (i.e., a uniform discrete measure on an appropriately chosen set of "equally spaced" points on a k-dimensional hypersphere) when both $d > 2$ and $k > 2$. So, we have

Corollary 9.1. There does not exist an S. H. optimal exact design δ^*_{SH} for S. H. regression of any order $d > 2$.

As promised in the previous section, we will now precisely specify the W. B. designs for S. H. and F. S. regression using the characterization given in Theorem 9.1.

In particular, it is fairly easy to see, if $(d+d') = 2d_o$ (using the notation of Section 8), that a WB design is a rotatable arrangement of order d_o; if $(d+d') = 2d_o + 1$, then a WB design can be either a rotatable arrangement of order (d_o+1) or a rotatable arrangement of order d_o with moments of order $(2d_o+1)$ equal to zero.

From the fact that we can fit an S. H. model of order $d = 2$ and an F. S. model of any order $d \geq 2$, but cannot fit a polynomial of degree $d > 1$, by using a rotatable arrangement of design points, we might say that, heuristically, the "loss of information" suffered by not using a rotatable design to fit the polynomial is measured, in some sense, by the "non-uniqueness" of the artificial polynomial response that can be constructed from the fitted S. H. or F. S. model using the techniques in Section 3.

One final comment is in order here before proceeding with the construction of optimal exact designs for S. H. and F. S. regression. Interestingly enough, Theorem 8.1 illustrates the only situation in the literature where a class of designs possesses so many optimal properties simultaneously. However, the results are not really as mystifying as they might seem at first glance if one takes note of Theorem 9.1 in conjunction with the geometrical interpretation of a rotatable arrangement and then remembers the implication of the general theory given in Section 6 (which essentially implies that a uniform weighting function dictates a uniform WB design) as well as the fact that designs constructed to minimize variance are, in general, "spread out" as much as possible over and extend to the limits of the region of interest.

Hoel [18] is the sole author in the literature to study F. S. and S. H. regression. He considers only G-optimality and characterizes the corresponding MV designs in terms of certain normalizing coefficients which are, in fact, the diagonal elements of $M^{-1}(\delta^*_{SH})$ and $M^{-1}(\delta^*_{FS})$. However, he does not suspect the valuable characterization of δ^*_{SH} and δ^*_{FS} emphasized in Theorem 9.1 nor does he mention the many optimal properties possessed by these designs. The response surface techniques developed for these functions in the next chapter are also new and quite useful.

If we write $M_{FS}(\alpha_1,\alpha_2) = \frac{1}{N}\sum_{i=1}^{N} (\cos\phi_i)^{\alpha_1} (\sin\phi_i)^{\alpha_2}$ and

$M_{SH}(\alpha_1, \alpha_2, \alpha_3) = \frac{1}{N}\sum_{i=1}^{N} (\cos\phi_i \sin\theta_i)^{\alpha_1} (\sin\phi_i \sin\theta_i)^{\alpha_2} (\cos\theta_i)^{\alpha_3}$,

then, by evaluating the necessary surface integrals on the unit circle and unit sphere, respectively, we are able to show that the moment requirements for δ^*_{FS} in terms of the points $\{\underset{\sim}{x}^{(i)}_{FS}\}_{i=1}^{N}$ and for δ^*_{SH} in terms of the points $\{\underset{\sim}{x}_{SH}^{(i)}\}_{i=1}^{N}$ are as follows:

$$(9.2) \qquad M_{FS}(\alpha_1,\alpha_2) = \Gamma\left(\frac{1+\alpha_1}{2}\right) \Gamma\left(\frac{1+\alpha_2}{2}\right) \Big/ \pi\, \Gamma\left(1 + \frac{\alpha_1+\alpha_2}{2}\right)$$

if all the α_j are even and is zero otherwise, where $0 \le \alpha_j \le 2d$, $j = 1, 2$, and $0 \le \alpha_1 + \alpha_2 \le 2d$;

$$(9.3) \qquad M_{SH}(\alpha_1,\alpha_2,\alpha_3) = \Gamma\left(\frac{1+\alpha_1}{2}\right) \Gamma\left(\frac{1+\alpha_2}{2}\right) \Gamma\left(\frac{1+\alpha_3}{2}\right) \Big/ 2\pi\, \Gamma\left(1 + \frac{1+\alpha_1+\alpha_2+\alpha_3}{2}\right)$$

if all the α_j are even and is zero otherwise, where $0 \le \alpha_j \le 2d$, $j = 1, 2, 3$, and $0 \le \alpha_1 + \alpha_2 + \alpha_3 \le 2d$.

Another way to verify the values of the non-zero moments above is to expand $\frac{1}{N}\sum_{i=1}^{N}(x_{1i}^2 + x_{2i}^2 + \ldots + x_{ki}^2)^j = 1$ sequentially for $j = 1, 2, \ldots, d$ using either $k = 2$ or 3, and, at each stage, to introduce the relationships between moments of the same order which are given by (9.1).

Now, consider an arrangement of p points equally spaced on the circumference of a circle. It has been shown in [5] that $p > 2d$ is sufficient for this arrangement to be rotatable of order d; the "necessity" argument can be found in [14]. Thus, after appealing to Theorem 9.1, we have the general result which now follows.

<u>THEOREM 9.2</u>. For F. S. regression of order d, an optimal <u>exact</u> design δ_{FS}^* assigns equal weight $1/p$ to each of the distinct points $\phi_i = 2\pi i/p$, $i = 1, 2, \ldots, p$, where $p > 2d$.

Note that it was not necessary to appeal to (9.2) in order to identify the optimal spacing given in the above theorem. On the other hand, we will need to use (9.3) in the construction of optimal <u>exact</u> designs for S. H. regression of orders one and two.

Now, adopting the notation of Bose and Draper in [1], we let the set of $p = 24$ points $G(a,b,c)$ be defined as

$$(9.4) \qquad G(a,b,c) = \{(\pm a,\pm b,\pm c),(\pm c,\pm a,\pm b),(\pm b,\pm c,\pm a)\},$$

where the restriction $a^2 + b^2 + c^2 = 1$ is required since, by Theorem 9.1, each point of $G(a,b,c)$ is to have the general representation $\underset{\sim}{x}_{SH}$. Without loss of generality, we can take $a \geq 0$, $b \geq 0$, $c \geq 0$. If θ and ϕ are the colatitude and azimuth, respectively, of a point on the surface of a unit sphere, then $\frac{1}{4}G(1,0,0)$ will represent the six vertices of an octahedron, $\frac{1}{3}G(1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})$ the eight vertices of a cube, and $\frac{1}{2}G(a,b,0)$ the twelve vertices of an icosahedron.

Now, from the form of (9.4), it is clear that the "odd moment" requirements of (9.3) are satisfied by the set of points $G(a,b,c)$, where $M_{SH}(\alpha_1,\alpha_2,\alpha_3)$ is an "odd moment" if at least one α_i is odd. Naturally, then, $M_{SH}(\alpha_1,\alpha_2,\alpha_3)$ is said to be an "even moment" if all α_i are even.

From (9.3), the even moment requirements for $d = 1$ are $M_{SH}(0,0,0) = 1$ (which is always satisfied) and $M_{SH}(2,0,0) = M_{SH}(0,2,0) = M_{SH}(0,0,2) = 1/3$. So, it is easy to see that $\frac{1}{4}G(1,0,0)$ and $\frac{1}{3}G(1/\sqrt{3},\ 1/\sqrt{3},\ 1/\sqrt{3})$ are S. H. optimal exact designs of order one. The four vertices of a tetrahedron (which is the G-optimal design of order one suggested by Hoel [18]) is yet another example of an S. H. optimal exact design of order one, but it is useless without augmentation except for purposes of estimation since there will be no degrees of freedom for error after fitting the first order S. H. model using this design.

For S. H. regression of order $d = 2$, the corresponding optimal exact designs must satisfy, in addition to the even moment requirements for d = 1 given earlier, the following relations: $M_{SH}(4,0,0) = M_{SH}(0,4,0) = M_{SH}(0,0,4) = 1/5$ and $M_{SH}(2,2,0) = M_{SH}(2,0,2) = M_{SH}(0,2,2) = 1/15$.

One can readily verify that the p = 20 vertices of the dodecahedron $\{\frac{1}{2}G(a,b,0) + \frac{1}{3}G(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}$, where $a^2 = \frac{1}{2}(1 + \sqrt{5}/3)$ and $b^2 = \frac{1}{2}(1 - \sqrt{5}/3)$, provide an exact S. H. optimal design of order two. The values given for a^2 and b^2 are the solutions of the two equations $a^2 + b^2 = 1$ and $a^4 + b^4 = 7/9$; it can be shown that these equations determine the values of a^2 and b^2 so that all the necessary even moment conditions are satisfied. Similarly, it is easy to show that the 28 point design $\{\frac{1}{2}G(a,b,0) + \frac{1}{3}G(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3}) + \frac{1}{3}G(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})\}$ is an S. H. optimal exact design of order two if $a^2 = \frac{1}{2}(1 + \frac{1}{3}\sqrt{41/5})$ and $b^2 = \frac{1}{2}(1 - \frac{1}{3}\sqrt{41/5})$. This design is an example of a <u>sequential</u> design in the sense that the two $\frac{1}{3}G(1/\sqrt{3}, 1/\sqrt{3}, 1/\sqrt{3})$ designs could initially be used together to fit a first-order S. H. model, and, if subsequently, a test of lack of fit (based on 4 degrees of freedom) was found to be significant, the remaining 12 points of $\frac{1}{2}G(a,b,0)$ (where the appropriate values of a^2 and b^2 are as given) could be added to provide the complete 28 point design of order two.

It is also possible to construct <u>infinite</u> classes of second order exact S. H. optimal designs. For example, consider the set of p = 24 points $\{\frac{1}{2}G(a_1,b_1,0) + \frac{1}{2}G(a_2,b_2,0)\}$. Values of a_1, b_1, a_2 and b_2 that satisfy the required moment conditions can be found by solving the set of simultaneous equations $a_1^2 + b_1^2 = 1$, $a_2^2 + b_2^2 = 1$, and $a_1^4 + b_1^4 + a_2^4 + b_2^4 = 6/5$. These equations generate a quadratic equation in a_1^2 of the form $a_1^4 - a_1^2 - a_2^2(1-a_2^2) + 2/5 = 0$, which has the two roots $1/2 \pm [a_2^2(1-a_2^2) - 3/20]^{1/2}$. Since $0 \leq a_1^2 \leq 1$, we require

$3/20 \leq a_2^2(1-a_2^2) \leq 2/5$.

The upper bound is obviously satisfied since a_2^2 must lie between zero and one, and so we need only have $a_2^4 - a_2^2 + 3/20 \leq 0$ which is satisfied if $(1/2 - 1/\sqrt{10}) \leq a_2^2 \leq (1/2 + 1/\sqrt{10})$. Thus, we have succeeded in constructing an infinite class of exact S. H. optimal designs of order two, each design in the class consisting of the 24 points $\{\frac{1}{2}G(a_1,b_1,0) + \frac{1}{2}G(a_2,b_2,0)\}$ and depending only on the single parameter a_2^2, where $(1/2 - 1/\sqrt{10}) \leq a_2^2 \leq (1/2 + 1/\sqrt{10})$, $b_2^2 = (1 - a_2^2)$, $a_1^2 = \frac{1}{2} + [a_2^2(1 - a_2^2) - 3/20]^{1/2}$, $b_1^2 = (1 - a_1^2)$.

As another example, consider the set of 24 points $G(a,b,c)$. Values of a, b and c that satisfy the appropriate moment requirements are found by solving the pair of equations $a^2 + b^2 + c^2 = 1$ and $a^4 + b^4 + c^4 = 3/5$, which together yield the quadratic in a^2 of the form $a^4 - (1 - c^2)\,a^2 - (c^2 - c^4 - 1/5) = 0$, whose two roots expressed as a function of c^2 are $\frac{1}{2}[(1-c^2) \pm (2c^2 - 3c^4 + 1/5)^{1/2}]$. Now, we require firstly that $(2c^2 - 3c^4 + 1/5) \geq 0$ which implies that $0 \leq c^2 \leq \frac{1}{3}(1 + 2\sqrt{2/5})$, secondly that $(1 - c^2) - (2c^2 - 3c^4 + 1/5)^{1/2} \geq 0$ which is equivalent to either $0 \leq c^2 \leq \frac{1}{2}(1 - 1/\sqrt{5})$ or $\frac{1}{2}(1 + 1/\sqrt{5}) \leq c^2 \leq 1$, and finally that $(1 - c^2) + (2c^2 - 3c^4 + 1/5)^{1/2} \leq 2$ which is satisfied if $c^2 \geq 0$. Combining these restrictions, we have another infinite class of exact S. H. optimal designs of order two, each design in this class consisting of the 24 points $G(a,b,c)$ and depending only on the one parameter c^2, where either $0 \leq c^2 \leq \frac{1}{2}(1 - 1/\sqrt{5})$ or $\frac{1}{2}(1 + 1/\sqrt{5}) \leq c^2 \leq \frac{1}{3}(1 + 2\sqrt{2/5})$, $a^2 = \frac{1}{2}[(1 - c^2) + (2c^2 - 3c^4 + 1/5)^{1/2}]$, and $b^2 = (1 - a^2 - c^2)$. In particular, the special cases $c^2 = 0$, $\frac{1}{2}(1 - 1/\sqrt{5})$ and $\frac{1}{2}(1 + 1/\sqrt{5})$ give the twelve vertices of an icosahedron. Other second-order exact S. H. optimal designs with as many points as desired can similarly be constructed. An example of an S. H. optimal approximate design of order three is given in [18].

For any of the previously constructed exact S. H. optimal designs, to convert a design point $\underset{\sim}{x}_{SH}^{(i)}$ to a point of the form (θ_i, ϕ_i), one would use (3.8). However, it is necessary to make this transformation to locate in terms of θ and ϕ the points at which responses are to be observed only when the region of interest is two-dimensional (see Section One), for, in this case, the elements of $\rho\underset{\sim}{x}_{SH}$ are artificial and do not represent the rectangular coordinates of a point on the surface of a sphere of radius ρ.

Also, since $\eta_d(\theta,\phi;\underset{\sim}{\beta}_{SH})$ is a function of θ and ϕ only through the elements of $\underset{\sim}{x}_{SH}$ (see Remark 3.1), we can write (1.11) for n = 1 and 2, respectively, as

$$(9.5) \qquad U_1(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(1)}) \equiv U_1(\underset{\sim}{x}_{SH};\underset{\sim}{\beta}_{SH}^{(1)}) = \alpha_{00} + \alpha_{10}x_3 + \alpha_{11}x_1 + \beta_{11}x_2,$$

and

$$(9.6) \qquad U_2(\theta,\phi;\underset{\sim}{\beta}_{SH}^{(2)}) \equiv U_2(\underset{\sim}{x}_{SH};\underset{\sim}{\beta}_{SH}^{(2)}) = \alpha_{20}\,\tfrac{1}{2}(3x_3^2 - 1) + \alpha_{21}(3x_1x_3)$$
$$+ \beta_{21}(3x_2x_3) + \alpha_{22}3(x_1^2 - x_2^2) + \beta_{22}(6x_1x_2).$$

So, computationally-wise, we are able to fit S. H. models of orders one and two by least squares using the exact S. H. optimal designs constructed earlier in this section without having to convert the points from the form $\underset{\sim}{x}_{SH}$ to the form (θ,ϕ) in order to be able to perform the necessary calculations.

Since we will be concerned in the next chapter with studying the properties of S. H. and F. S. models which have been fitted by least squares hopefully employing the optimal exact designs presented earlier in this section, let us discuss the methodology associated with the use of these designs. In particular, let

$$\underset{\sim}{b}'_{FS} = (\underset{\sim}{b}^{(0)'}_{FS}, \underset{\sim}{b}^{(1)'}_{FS}, \ldots, \underset{\sim}{b}^{(d)'}_{FS}) \tag{9.7}$$

and

$$\underset{\sim}{b}'_{SH} = (\underset{\sim}{b}^{(0)'}_{SH}, \underset{\sim}{b}^{(1)'}_{SH}, \ldots, \underset{\sim}{b}^{(d)'}_{SH}) \tag{9.8}$$

denote the vectors of least squares estimates of the elements of $\underset{\sim}{\beta}_{FS}$ and $\underset{\sim}{\beta}_{SH}$, respectively, where

$$\underset{\sim}{b}^{(0)}_{FS} = a_0,\ \underset{\sim}{b}^{(n)'}_{FS} = (a_n, b_n),\ n = 1, 2, \ldots, d, \tag{9.9}$$

and

$$\underset{\sim}{b}^{(0)}_{SH} = a_{00},\ \underset{\sim}{b}^{(n)'}_{SH} = (a_{n0}, a_{n1}, b_{n1}, \ldots, a_{nn}, b_{nn}),\ n = 1, 2, \ldots, d. \tag{9.10}$$

If we use the optimal spacing advocated in Theorem 9.2 to fit an F. S. model of order d, then it is easy to see that

$$a_0 = \frac{1}{N} \sum_{i=1}^{p} \sum_{j=1}^{r} y_j(2\pi i/p) = \bar{y}, \tag{9.11}$$

$$a_n = \frac{2}{p} \sum_{i=1}^{p} \cos(2n\pi i/p)\ \bar{y}(2\pi i/p),\ n = 1, 2, \ldots, d, \tag{9.12}$$

$$b_n = \frac{2}{p} \sum_{i=1}^{p} \sin(2n\pi i/p)\ \bar{y}(2\pi i/p),\ n = 1, 2, \ldots, d, \tag{9.13}$$

where $y_j(2\pi i/p)$ is the observed response for the jth replication of the ith point $\phi_i = 2\pi i/p$, $j = 1, 2, \ldots, r$, $i = 1, 2, \ldots, p$,

$\bar{y}(2\pi i/p) = \frac{1}{r}\sum_{j=1}^{r} y_j(2\pi i/p)$, and $N = pr$. Also, it follows from what has been done that $\mathrm{Var}(\underset{\sim}{b}_{FS}) = D_{FS}^{-1}\,\sigma^2/N$ and that $\mathrm{Var}\ \hat{y}_d(\phi;\underset{\sim}{b}_{FS}) = (2d+1)\ \sigma^2/N$.

Similarly, for S. H. regression of order two, we use the computational forms (9.5) and (9.6) to write

$$a_{00} = \frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}) = \bar{y},$$

$$a_{10} = \frac{3}{N}\sum_{i=1}^{p} x_{3i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}),$$

$$a_{11} = \frac{3}{N}\sum_{i=1}^{p} x_{1i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}),$$

$$b_{11} = \frac{3}{N}\sum_{i=1}^{p} x_{2i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}),$$

$$a_{20} = \frac{15}{2N}\sum_{i=1}^{p} x_{3i}^2\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}) - \frac{5}{2}a_{00},$$

$$a_{21} = \frac{5}{N}\sum_{i=1}^{p} x_{1i}x_{3i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)})\ ,$$

$$b_{21} = \frac{5}{N}\sum_{i=1}^{p} x_{2i}x_{3i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}),$$

$$a_{22} = \frac{5}{2N}\sum_{i=1}^{p} x_{1i}^2\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}) + \frac{1}{6}(a_{20} - 5a_{00}),$$

$$b_{22} = \frac{5}{2N}\sum_{i=1}^{p} x_{1i}x_{2i}\sum_{j=1}^{r_i} y_j(\underset{\sim}{x}_{SH}^{(i)}),$$

where $y_j(\underset{\sim}{x}_{SH}^{(i)})$ is the observed response for the jth replication of the ith point $\underset{\sim}{x}_{SH}^{(i)'} = (x_{3i}, x_{1i}, x_{2i}) = (\cos\theta_i, \cos\phi_i\ \sin\theta_i, \sin\phi_i\ \sin\theta_i)$, $j = 1, 2, \ldots, r_i$, $i = 1, 2, \ldots, p$, and $N = \sum_{i=1}^{p} r_i$. Also it follows that $\mathrm{Var}(\underset{\sim}{b}_{SH}) = D_{SH}^{-1}\,\sigma^2/N = G_{SH}^{1/2}\,D_{SH}^{*-1}\,G_{SH}^{1/2}\,\sigma^2/N$ and that $\mathrm{Var}\ \hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH}) = (d+1)^2\,\sigma^2/N$. In particular, for $d = 2$, we have

(9.14) $\mathrm{Var}(\underset{\sim}{b}_{SH}) = \mathrm{diag}(1;\ 3,3,3;\ 5,5/3,5/3,5/12,5/12)\ \sigma^2/N.$

In the above framework, it is easy to form the analysis of variance tables associated with these designs.

ANOVA TABLE FOR F. S. REGRESSION OF ORDER d USING δ^*_{FS}

Source	D.F.	Sum of Squares
a_0	1	$pr\ a_0^2$
a_1	1	$pr\ a_1^2/2$
b_1	1	$pr\ b_1^2/2$
a_2	1	$pr\ a_2^2/2$
b_2	1	$pr\ b_2^2/2$
$\vdots$	$\vdots$	$\vdots$
a_d	1	$pr\ a_d^2/2$
b_d	1	$pr\ b_d^2/2$
Lack of fit	$p-(2d+1)$	By difference
Pure error	$p(r-1)$	$\sum_{i=1}^p \sum_{j=1}^r (y_j(2\pi i/p)-\bar{y}(2\pi i/p))^2$
Total	pr	$\sum_{i=1}^p \sum_{j=1}^r y_j^2(2\pi i/p)$

ANOVA TABLE FOR S. H. REGRESSION OF ORDER TWO USING δ^*_{SH}

Source	D.F.	Sum of Squares
a_{00}	1	Na^2_{00}
a_{10}	1	$Na^2_{10}/3$
a_{11}	1	$Na^2_{11}/3$
b_{11}	1	$Nb^2_{11}/3$
a_{20}	1	$Na^2_{20}/5$
a_{21}	1	$3Na^2_{21}/5$
b_{21}	1	$3Nb^2_{21}/5$
a_{22}	1	$12Na^2_{22}/5$
b_{22}	1	$12Nb^2_{22}/5$
Lack of fit	$(p-9)$	By difference
Pure error	$(N-p)$	$\sum_{i=1}^{p} \sum_{j=1}^{r_i} (y_j(\underset{\sim}{x}^{(i)}_{SH}) - \bar{y}(\underset{\sim}{x}^{(i)}_{SH}))^2$
Total	$N=\sum_{i=1}^{p} r_i$	$\sum_{i=1}^{p} \sum_{j=1}^{r_i} y_j^2(\underset{\sim}{x}^{(i)}_{SH})$

Finally, as promised in Section 1, we will make some brief comments here concerning the use of direct products of F. S. regression functions as models for response functions which are known to be periodic in more than one variable. In general, to graduate a response function $\eta(x_1, x_2, \ldots, x_k; \underset{\sim}{\beta})$ on the k-dimensional hyperplane which is periodic in $x_1, x_2, \ldots, x_k$ for $k \geq 2$, we can take as a model the expression $\eta(\phi_1, \phi_2, \ldots, \phi_k; \underset{\sim}{\beta}) = \eta_{d_1}(\phi_1; \underset{\sim}{\beta}_{FS})\, \eta_{d_2}(\phi_2; \underset{\sim}{\beta}_{FS}) \cdots \eta_{d_k}(\phi_k; \underset{\sim}{\beta}_{FS})$, where $0 \leq \phi_i \leq 2\pi$ for $i = 1, 2, \ldots, k$, which can be written as the sum of $(2d_1+1)(2d_2+1) \cdots (2d_k+1)$ distinct terms. We can appeal to Theorem 9.2 of this section and to Theorem 2, p. 1099, of [18] to conclude that an MV (or G-optimal) exact design for fitting $\eta(\phi_1, \ldots, \phi_k; \underset{\sim}{\beta})$ by least squares is one that assigns equal weight to each of the $p_1 p_2 \cdots p_k$ points $(\phi_1, \phi_2, \ldots, \phi_k)$, where ϕ_i is allowed to run over the set of values $2\pi j/p_i$, $j = 1, 2, \ldots, p_i$, $p_i > 2d_i$ for every i, $i = 1, 2, \ldots, k$. These notions have not been pursued further.

CHAPTER IV

RESPONSE SURFACE TECHNIQUES FOR F. S. AND S. H. REGRESSION

10. Analysis of Fitted Second Order F. S. and S. H. Models

As stated in the introduction, the objective of any response surface study is to approximate as accurately as possible using some graduating function the form of an unknown response surface on some region of interest $\mathcal{X}$. One is then interested in using this graduating function to estimate the locations in $\mathcal{X}$ at which the surface has a maximum, minimum, or saddle point so that, for example, optimal settings of the factor levels can be specified. (Henceforth, such locations will be referred to as "stationary points".) Techniques for studying polynomials of degree two in this light were developed by Box and Wilson [6].

In this section, we present a general method for locating the stationary points of fitted F. S. and S. H. models of order two. Of course, it is suggested that the least squares fitting of these models be done using the designs presented in the previous section.

By noting that $\hat{y}_d(\phi;\underset{\sim}{b}_{FS})$ is a function of ϕ only through the elements of $\underset{\sim}{x}_{FS}$ and that $\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})$ is a function of θ and ϕ only through the elements of $\underset{\sim}{x}_{SH}$ (again, see Remark 3.1), we can conclude that the determination of the stationary points of fitted F. S. and S. H. models could be made by differentiating these functions with respect to the elements of $\underset{\sim}{x}_{FS}$ and $\underset{\sim}{x}_{SH}$, respectively, subject to the restrictions $\underset{\sim}{x}'_{FS}\underset{\sim}{x}_{FS} = 1$ and $\underset{\sim}{x}'_{SH}\underset{\sim}{x}_{SH} = 1$, rather than by performing the differentiations directly with respect to ϕ for the fitted F. S. model and with respect to θ and ϕ for the fitted S. H. model. As a matter of fact, an attempt to determine the stationary points by solving $\partial\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})/\partial\theta = 0$ and $\partial\hat{y}_d(\theta,\phi;\underset{\sim}{b}_{SH})/\partial\phi = 0$ in the case of S. H. regression or by solving $d\hat{y}_d(\phi;\underset{\sim}{b}_{FS})/d\phi = 0$ in the case of F. S. regression will not meet with any noticeable success, except possibly for the case $d = 1$. On the other hand, the technique we propose, which involves working with fitted F. S. and S. H. models of order two in polynomial form, is quite easy to use and yields results which lend themselves to direct interpretation. The relations (3.6) and (3.8) can then be used to convert stationary points $\underset{\sim}{x}_{FS}^{(o)}$ and $\underset{\sim}{x}_{SH}^{(o)}$ to points of the form ϕ_o and (θ_o,ϕ_o), respectively.

Now, using (3.7), we can write a fitted S. H. model of order two in the form

$$(10.1)\qquad \hat{y}_2(\underset{\sim}{x}_{SH};\underset{\sim}{b}_{SH}) = (a_{00}-a_{20}/2) + \underset{\sim}{x}'_{SH}\,\underset{\sim}{b}_{SH}^{(1)} + \underset{\sim}{x}'_{SH}\,B_{SH}\,\underset{\sim}{x}_{SH},$$

where $\underset{\sim}{x}_{SH}$ is given by (3.4), $\underset{\sim}{b}_{SH}^{(1)}$ by (9.10), and where

$$(10.2)\qquad B_{SH} = \frac{3}{2}\begin{bmatrix} a_{20} & a_{21} & b_{21} \\ a_{21} & 2a_{22} & 2b_{22} \\ b_{21} & 2b_{22} & -2a_{22} \end{bmatrix}.$$

Similarly, we may write (3.5) as

$$(10.3)\qquad \hat{y}_2(\underset{\sim}{x}_{FS};\underset{\sim}{b}_{FS}) = a_0 + \underset{\sim}{x}'_{FS}\,\underset{\sim}{b}_{FS}^{(1)} + \underset{\sim}{x}'_{FS}\,B_{FS}\,\underset{\sim}{x}_{FS},$$

where $\underset{\sim}{x}_{FS}$ is given by (3.3), $\underset{\sim}{b}_{FS}^{(1)}$ by (9.9), and where

$$(10.4) \qquad B_{FS} = \begin{bmatrix} a_2 & b_2 \\ b_2 & -a_2 \end{bmatrix} .$$

It will be shown later (Lemma 10.1) that the validity of the proposed techniques will not be affected by the non-uniqueness of the polynomial representations (10.1) and (10.3).

Now, from the above, it follows that, in general, we are interested in determining the stationary points of the function

$$(10.5) \qquad f(\underset{\sim}{x}) = b_0 + \underset{\sim}{x}'\,\underset{\sim}{b} + \underset{\sim}{x}'\,B\,\underset{\sim}{x},$$

subject to the restriction $\underset{\sim}{x}'\,\underset{\sim}{x} = 1$, where b_0 is a known real constant and where $\underset{\sim}{b}$ and B are, respectively, a (k x 1) vector and a (k x k) symmetric matrix of known real constants. This general description covers the situation of interest to us, namely the case when b_0, $\underset{\sim}{b}$ and B are stochastic and (10.5) is a specific realization of the random variable $f(\underset{\sim}{x})$; in particular, the work in Section 11 is based solely on this notion.

So, let $Q(\underset{\sim}{x}) = b_0 + \underset{\sim}{x}'\underset{\sim}{b} + \underset{\sim}{x}'B\underset{\sim}{x} - \mu(\underset{\sim}{x}'\underset{\sim}{x} - 1)$, where μ is a Lagrange multiplier. Then, it is a straightforward exercise to verify that a vector $\underset{\sim}{x}_0$ which satisfies the equation $dQ(\underset{\sim}{x})/d\underset{\sim}{x} = \underset{\sim}{0}$ is a solution of the pair of equations

$$(10.6) \qquad C\,\underset{\sim}{x} = -\frac{1}{2}\,\underset{\sim}{b} \quad \text{and} \quad \underset{\sim}{x}'\underset{\sim}{x} = 1,$$

where

$$(10.7) \qquad C = (B - \mu I_k).$$

Clearly, a different solution x_o of (10.6) will be obtained for each different value assigned to μ. Keeping μ arbitrary in (10.6) for the moment and remembering that $(B-\mu I_k)^{-1} = \mathrm{adj}(B-\mu I_k)/|B-\mu I_k|$, where $\mathrm{adj}(B-\mu I_k)$ is the transpose of the matrix of cofactors of $(B-\mu I_k)$, we see that the equation determining the permissible set of values of μ is

$$(10.8) \qquad b'[\mathrm{adj}(B-\mu I_k)]^2 b = 4|B-\mu I_k|^2 = 4(\mu-\lambda_1)^2(\mu-\lambda_2)^2\cdots(\mu-\lambda_k)^2,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_k$ are the k characteristic roots of B. Note that some of the λ's may be zero since B has not been assumed to be of full rank here. The second equality in (10.8) follows directly by pre- and post-multiplying $|B-\mu I_k|^2$ by $|P|^2$, where P is a real orthogonal matrix such that $PBP' = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$. (We will refer to this particular matrix P again.)

Now, (10.8) can equivalently be written as a polynomial of degree 2k in μ; moreover, this polynomial will have at least two real roots since (10.5) is a continuous real function on the compact set $\mathcal{X} = \{x: x'x = 1\}$ and hence there must exist points in $\mathcal{X}$ at which (10.5) achieves its maximum and its minimum. If a real root, μ_o say, of (10.8) is not a characteristic root of B as well, then the solution x_o of (10.6) corresponding to μ_o is simply $x_o = -\frac{1}{2}(B-\mu_o I_k)^{-1}b$. On the other hand, if $b'[\mathrm{adj}(B-\lambda_j I_k)]^2 b$ is zero for some λ_j, $j = 1,\ldots,k$, then λ_j is a root of multiplicity two of (10.8), $|B-\lambda_j I_k| = 0$, and the two equations $(B-\lambda_j I_k)x = -\frac{1}{2}\, b$ and $x'x = 1$ must be solved simultaneously.

<u>Remark 10.1.</u> It will be shown later in this section that the solution

of (10.6) using the largest real root of (10.8) is that value of $\underset{\sim}{x}$ which maximizes (10.5) subject to the constraint $\underset{\sim}{x}'\underset{\sim}{x} = 1$, and the solution of (10.6) using the smallest real root of (10.8) is that value of $\underset{\sim}{x}$ which minimizes (10.5) subject to $\underset{\sim}{x}'\underset{\sim}{x} = 1$.

Now, if we set $\underset{\sim}{b} = \underset{\sim}{b}_{FS}^{(1)}$ and $B = B_{FS}$ in (10.8), the resulting expression can be written in the form

$$
\begin{aligned}
(10.9)\quad & \mu^4 - \tfrac{1}{4}[(a_1^2 + b_1^2) + 8(a_2^2 + b_2^2)]\mu^2 + \tfrac{1}{2}[a_2(b_1^2 - a_1^2) - 2a_1b_1b_2]\mu \\
& + \tfrac{1}{4}(a_2^2 + b_2^2)[4(a_2^2 + b_2^2) - (a_1^2 + b_1^2)] = 0.
\end{aligned}
$$

Similarly, if we set $\underset{\sim}{b} = \underset{\sim}{b}_{SH}^{(1)}$ and $B = B_{SH}$ in (10.8), then a considerable amount of algebra gives

$$
\begin{aligned}
(10.10)\quad & \mu^6 - 3a_{20}\mu^5 + \tfrac{1}{4}\{9[a_{20}^2 - 2(a_{21}^2 + b_{21}^2) - 8(a_{22}^2 + b_{22}^2)] - (a_{10}^2 \\
& + a_{11}^2 + b_{11}^2)\}\mu^4 + \tfrac{3}{4}\{9[a_{20}(a_{21}^2 + b_{21}^2) + 8a_{20}(a_{22}^2 + b_{22}^2) \\
& - 2a_{21}(a_{21}a_{22} + b_{21}b_{22}) - 2b_{21}(a_{21}b_{22} - a_{22}b_{21})] + [a_{11}^2(a_{20} \\
& - 2a_{22}) + b_{11}^2(a_{20} + 2a_{22})] - 2[a_{10}(a_{11}a_{21} + b_{11}b_{21}) \\
& + 2a_{11}b_{11}b_{22}]\}\mu^3 + \tfrac{9}{16}\{9[(a_{21}^2 + b_{21}^2) + 4(a_{22}^2 + b_{22}^2)]^2 \\
& + 36a_{20}[a_{21}(a_{21}a_{22} + b_{21}b_{22}) + b_{21}(a_{21}b_{22} - a_{22}b_{21}) \\
& - 2a_{20}(a_{22}^2 + b_{22}^2)] + 2a_{10}a_{20}(a_{11}a_{21} + b_{11}b_{21}) \\
& - 12a_{10}[a_{11}(a_{21}a_{22} + b_{21}b_{22}) + b_{11}(a_{21}b_{22} - a_{22}b_{21})] \\
& - a_{10}^2[(a_{21}^2 + b_{21}^2) - 8(a_{22}^2 + b_{22}^2)] - (a_{11}^2 + b_{11}^2)[a_{20}^2 \\
& + 4(a_{22}^2 + b_{22}^2)] + 2(a_{11}b_{21} - a_{21}b_{11})^2 - (a_{11}a_{21} + b_{11}b_{21})^2 \\
& + 8a_{20}[a_{11}(a_{11}a_{22} + b_{11}b_{22}) + b_{11}(a_{11}b_{22} - a_{22}b_{11})]\}\mu^2 \\
& + \tfrac{27}{16}\{18[(a_{21}^2 + b_{21}^2) + 4(a_{22}^2 + b_{22}^2)][a_{21}(a_{21}a_{22} + b_{21}b_{22}) \\
& + b_{21}(a_{21}b_{22} - a_{22}b_{21}) - 2a_{20}(a_{22}^2 + b_{22}^2)] - a_{11}^2(a_{20} \\
& - 2a_{22})(b_{21}^2 + 2a_{20}a_{22}) - b_{11}^2(a_{20} + 2a_{22})(a_{21}^2 - 2a_{20}a_{22}) \\
& + 2(a_{21}b_{21} - 2a_{20}b_{22})[a_{20}a_{11}b_{11} - b_{22}(a_{11}^2 + b_{11}^2)]
\end{aligned}
$$

$$+ 2(a_{21}a_{22} + b_{21}b_{22})[2a_{10}a_{11}a_{20} - a_{21}(a_{10}^2 + a_{11}^2)]$$
$$+ 2(a_{21}b_{22} - a_{22}b_{21})[2a_{10}a_{20}b_{11} - b_{21}(a_{10}^2 + b_{11}^2)]\}\mu$$
$$+ \frac{81}{64}\{36[a_{21}(a_{21}a_{22} + b_{21}b_{22}) + b_{21}(a_{21}b_{22} - a_{22}b_{21})$$
$$- 2a_{20}(a_{22}^2 + b_{22}^2)]^2 - a_{11}^2(b_{21}^2 + 2a_{20}a_{22})^2$$
$$- b_{11}^2(a_{21}^2 - 2a_{20}a_{22})^2 - 16a_{10}^2(a_{22}^2 + b_{22}^2)^2 - 4(a_{21}a_{22}$$
$$+ b_{21}b_{22})[a_{10}b_{11}(a_{21}b_{21} - 2a_{20}b_{22}) + (a_{10}^2 + a_{11}^2)(a_{21}a_{22}$$
$$+ b_{21}b_{22}) + 2a_{11}b_{11}(a_{21}b_{22} - a_{22}b_{21}) - a_{10}a_{11}(b_{21}^2 + 4b_{22}^2)$$
$$- 2a_{10}a_{11}a_{22}(a_{20} + 2a_{22})] - [a_{11}(a_{21}b_{21} - 2a_{20}b_{22})$$
$$+ 2a_{10}(a_{21}b_{22} - a_{22}b_{21})]^2 + 4b_{11}(a_{21}b_{22} - a_{22}b_{21})[a_{10}(a_{21}^2$$
$$+ 4b_{22}^2) - 2a_{10}a_{22}(a_{20} - 2a_{22}) - b_{11}(a_{21}b_{22} - a_{22}b_{21})]$$
$$+ b_{11}(a_{21}b_{21} - 2a_{20}b_{22})[2a_{11}(a_{21}^2 + b_{21}^2) - b_{11}(a_{21}b_{21}$$
$$- 2a_{20}b_{22})]\} = 0.$$

The determination of the values of the coefficients in (10.9) and especially in (10.10) is best and most easily done by writing a simple computer program to perform the necessary arithmetical manipulations. This method of approach attempts to avoid possible round-off errors which seriously affect the values of the roots of these polynomials (which are also easily obtained by standard computer techniques), thus altering the corresponding stationary points found by solving the equations (10.6). Since it is reasonable to assume that the random matrices $(B_{FS}-\mu_o I_2)$ and $(B_{SH}-\mu_o I_3)$ will be non-singular with probability one, a check on the accuracy of the calculation of the coefficients and roots of (10.9) and (10.10) is to see how close $\underset{\sim}{x}_o' \underset{\sim}{x}_o$ is to one, where, in our general notation, $\underset{\sim}{x}_o = -\frac{1}{2}(B-\mu_o I_k)^{-1}\underset{\sim}{b}$. (The case when $|B-\mu_o I_k|$ is close to zero will be discussed in the next section.)

The following lemma removes the doubt raised by the non-uniqueness of the representations (10.1) and (10.3).

Lemma 10.1. The method proposed for determining the set of stationary points of a fitted F. S. or S. H. model of order two will generate the correct set of points regardless of the choice of the second degree polynomial representation of $\hat{y}_2(\phi;\underset{\sim}{b}_{FS})$ or $\hat{y}_2(\theta,\phi;\underset{\sim}{b}_{SH})$ used to employ (10.6) and (10.8).

Proof. Let $f_1(\underset{\sim}{x}) = b_1 + \underset{\sim}{x}'\underset{\sim}{b}_1 + \underset{\sim}{x}'B_1\underset{\sim}{x}$ and $f_2(\underset{\sim}{x}) = b_2 + \underset{\sim}{x}'\underset{\sim}{b}_2 + \underset{\sim}{x}'B_2\underset{\sim}{x}$ denote two different second degree polynomial representations for either $\hat{y}_2(\phi;\underset{\sim}{b}_{FS})$ or $\hat{y}_2(\theta,\phi;\underset{\sim}{b}_{SH})$, where $\underset{\sim}{x}'\underset{\sim}{x} = 1$. It is clear from the way in which these representations are formed that the relations

$$(10.11) \qquad b_2 = b_1 + c,\ \underset{\sim}{b}_1 = \underset{\sim}{b}_2,\ B_2 = B_1 - cI_k$$

hold, where c is a non-zero real constant. So, using (10.11), we can write the equations (10.6) and (10.8) associated with the polynomial representation $f_2(\underset{\sim}{x})$ in the form $[B_1-(\mu+c)I_k]\underset{\sim}{x} = -\frac{1}{2}\underset{\sim}{b}_1$, $\underset{\sim}{x}'\underset{\sim}{x} = 1$, $\underset{\sim}{b}_1'\{\mathrm{adj}[B_1-(\mu+c)I_k]\}^2\underset{\sim}{b}_1 = 4|B_1-(\mu+c)I_k|^2$, and these are easily seen to be equivalent to (10.6) and (10.8) under the representation $f_1(\underset{\sim}{x})$ with $(\mu+c)$ replaced by μ^*, say. Thus, the set of solutions, $\{\underset{\sim}{x}_o^{(1)}\}$ say, obtained using the representation $f_1(\underset{\sim}{x})$ is identical to the set of solutions $\{\underset{\sim}{x}_o^{(2)}\}$ obtained using $f_2(\underset{\sim}{x})$. This completes the proof.

In particular, it follows from the above argument and from the form of (10.6) that the same set of matrices $\{C\}$, where C is given by (10.7), will be obtained regardless of the choice of the polynomial representation. Briefly, what is happening here is that a change in the polynomial representation affects both the form of B and the set of roots of (10.8) in such a way that the set of matrices $\{C\}$, the vector $\underset{\sim}{b}$, and hence the set of stationary points $\{\underset{\sim}{x}_o\}$ all remain invariant.

Now, suppose $\underset{\sim}{x}_o$ is a solution of the equations $C_o\underset{\sim}{x} = -\frac{1}{2}\underset{\sim}{b}$ and $\underset{\sim}{x}'\underset{\sim}{x} = 1$, where $C_o = (B-\mu_o I_k)$ and where μ_o is a real root of (10.8). Then, since $\underset{\sim}{x}'\underset{\sim}{x} = \underset{\sim}{x}_o'\underset{\sim}{x}_o = 1$, we have

$$\begin{aligned} f(\underset{\sim}{x}) - f(\underset{\sim}{x}_o) &= (b_0 + \underset{\sim}{x}'\underset{\sim}{b} + \underset{\sim}{x}'B\underset{\sim}{x}) - (b_0 + \underset{\sim}{x}_o'\underset{\sim}{b} + \underset{\sim}{x}_o'B\underset{\sim}{x}_o) \\ &= (\underset{\sim}{x}-\underset{\sim}{x}_o)'\underset{\sim}{b} + \underset{\sim}{x}'C_o\underset{\sim}{x} - \underset{\sim}{x}_o'C_o\underset{\sim}{x}_o \\ &= (\underset{\sim}{x}-\underset{\sim}{x}_o)'\underset{\sim}{b} + 2(\underset{\sim}{x}-\underset{\sim}{x}_o)'C_o\underset{\sim}{x}_o + (\underset{\sim}{x}-\underset{\sim}{x}_o)'C_o(\underset{\sim}{x}-\underset{\sim}{x}_o) \\ &= (\underset{\sim}{x}-\underset{\sim}{x}_o)'C_o(\underset{\sim}{x}-\underset{\sim}{x}_o). \end{aligned}$$

If the orthogonal matrix P is chosen as before, then it is clear that we can write

$$f(\underset{\sim}{x}) - f(\underset{\sim}{x}_o) = (\underset{\sim}{x}-\underset{\sim}{x}_o)'P'[\mathrm{diag}(\lambda_1-\mu_o,\ \lambda_2-\mu_o,\ldots,\ \lambda_k-\mu_o)]P(\underset{\sim}{x}-\underset{\sim}{x}_o), \text{ or}$$

$$(10.12)\qquad f(\underset{\sim}{x}) - f(\underset{\sim}{x}_o) = \textstyle\sum_{i=1}^{k}(\lambda_i-\mu_o)w_i^2,\ (w_1,\ w_2,\ldots,\ w_k) = (\underset{\sim}{x}-\underset{\sim}{x}_o)'P'.$$

Geometrically, the expression (10.12) can be considered to be the result of shifting the origin $\underset{\sim}{0}$ of the $\underset{\sim}{x}$'s to the stationary point $\underset{\sim}{x}_o$ and of rotating the coordinate axes in terms of $x_1, x_2,\ldots, x_k$ so that they become axes expressed in terms of $w_1, w_2,\ldots, w_k$. It is easy to see that the quantities $(\lambda_1-\mu_o), (\lambda_2-\mu_o),\ldots, (\lambda_k-\mu_o)$ are the characteristic roots of $C_o = (B-\mu_o I_k)$.

Note that the expression (10.12) illustrates the change in response on moving away from the stationary point $\underset{\sim}{x}_o$. In particular, if the characteristic roots of C_o are all non-negative (non-positive), then (10.5) takes on its minimum (maximum) value at $\underset{\sim}{x}_o$ in the set $\mathcal{X} = \{\underset{\sim}{x}\colon \underset{\sim}{x}'\underset{\sim}{x} = 1\}$. If at least two of the characteristic roots of C_o are of different sign, then $\underset{\sim}{x}_o$ is said to be a saddle point. It is important to realize that a stationary point $\underset{\sim}{x}_o$ (whether it be a maximum, minimum, or saddle point) will be unique only when $|C_o| \neq 0$.

Since the maximum (minimum) of (10.5) subject to $\underset{\sim}{x}'\underset{\sim}{x} = 1$ must be found using this procedure, it follows that there will be a μ_o greater (smaller) than or equal to all the λ's, and this μ_o will obviously be the largest (smallest) root of (10.8). In a given situation, then, one would really need to solve (10.6) using only the largest and smallest roots of (10.8) since the extremum points (and not the saddle points) are of main interest. One can now appreciate the import of Remark 10.1.

These techniques can be used for all the regions on which we have considered these F. S. and S. H. functions to be defined, although the transformations, etc., actually have a geometrical interpretation when the regions of interest are the circumference of a circle and the surface of a sphere.

For completeness, let us characterize the extremum points $\underset{\sim}{x}_{max}$ and $\underset{\sim}{x}_{min}$ for the following special cases. If $\underset{\sim}{b} \neq 0$ and $B = 0$, then $\underset{\sim}{x}_{max} = \underset{\sim}{b}/(\underset{\sim}{b}'\underset{\sim}{b})^{1/2}$ and $\underset{\sim}{x}_{min} = -\underset{\sim}{b}/(\underset{\sim}{b}'\underset{\sim}{b})^{1/2}$, where we agree to take the positive square root of $\underset{\sim}{b}'\underset{\sim}{b}$. If $\underset{\sim}{b} = 0$ and $B \neq 0$, then $\underset{\sim}{x}_{max}$ and $\underset{\sim}{x}_{min}$ are the characteristic vectors of unit length associated with the largest and smallest characteristic roots of B, respectively.

11. Confidence Regions for Stationary Points of Second Order F. S. and S. H. Models

If $\underset{\sim}{x}_o$ denotes a solution of the equations (10.6) corresponding to a real root μ_o of (10.8), then it follows, ignoring bias effects, that the true (but unknown) value, $\underset{\sim}{x}_o^*$, of this stationary point must satisfy the matrix equation $H_{\underset{\sim}{x}_o^*}$: $\underset{\sim}{\delta}_{\underset{\sim}{x}_o^*} = \frac{1}{2}E(\underset{\sim}{b}) + E(C_o)\underset{\sim}{x}_o^* = \underset{\sim}{0}$, where $\underset{\sim}{x}_o^{*\prime}\underset{\sim}{x}_o^* = 1$. Clearly, $H_{\underset{\sim}{x}_o^*}$ is in the form of a general linear hypothesis and any procedure for testing $H_{\underset{\sim}{x}}$: $\underset{\sim}{\delta}_{\underset{\sim}{x}} = \frac{1}{2}E(\underset{\sim}{b}) + E(C_o)\underset{\sim}{x} = \underset{\sim}{0}$ for any given $\underset{\sim}{x}$ (in our case, for any $\underset{\sim}{x}$ for which $\underset{\sim}{x}'\underset{\sim}{x} = 1$) can be used to construct a confidence region for $\underset{\sim}{x}_o^*$. (For general discussions along this line, see [4] and [24]. An application related to rotatable designs is given in [5].)

Now, let $\underset{\sim}{d}_{\underset{\sim}{x}} = \frac{1}{2}\underset{\sim}{b} + C_o\underset{\sim}{x}$ be the estimate of $\underset{\sim}{\delta}_{\underset{\sim}{x}}$. Then, it is easy to see that the elements of $\underset{\sim}{d}_{\underset{\sim}{x}}$ are linear combinations of the elements of $\underset{\sim}{b}$ and B if one notes that $\mu_o = \frac{1}{2}\underset{\sim}{x}'\underset{\sim}{b} + \underset{\sim}{x}'B\underset{\sim}{x}$. This fact enables us to write down the dispersion matrices $V(\underset{\sim}{d}_{\underset{\sim}{x}_{FS}})$ and $V(\underset{\sim}{d}_{\underset{\sim}{x}_{SH}})$ of $\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}$ and $\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$ corresponding to the special cases $\underset{\sim}{x} = \underset{\sim}{x}_{FS}$, $\underset{\sim}{b} = \underset{\sim}{b}_{FS}^{(1)}$, $B = B_{FS}$ and $\underset{\sim}{x} = \underset{\sim}{x}_{SH}$, $\underset{\sim}{b} = \underset{\sim}{b}_{SH}^{(1)}$, $B = B_{SH}$, respectively. In particular, using (9.14) and the fact that $Var(\underset{\sim}{b}_{FS}) = D_{FS}^{-1}\sigma^2/N$ and also employing the relations $\underset{\sim}{x}_{FS}'\underset{\sim}{x}_{FS} = 1$ and $\underset{\sim}{x}_{SH}'\underset{\sim}{x}_{SH} = 1$ at the proper times, we can show after some labor that $V(\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}) = \frac{5}{2}(I_2 - \underset{\sim}{x}_{FS}\underset{\sim}{x}_{FS}')\sigma^2/N$ and $V(\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}) = \frac{9}{2}(I_3 - \underset{\sim}{x}_{SH}\underset{\sim}{x}_{SH}')\sigma^2/N$. It is important to realize that the matrices $(I_2 - \underset{\sim}{x}_{FS}\underset{\sim}{x}_{FS}')$ and $(I_3 - \underset{\sim}{x}_{SH}\underset{\sim}{x}_{SH}')$ are both idempotent and are of ranks one and two, respectively. The reason that the dispersion matrices $V(\underset{\sim}{d}_{\underset{\sim}{x}_{FS}})$ and $V(\underset{\sim}{d}_{\underset{\sim}{x}_{SH}})$ turn out to be singular is due to the fact that the general relation $\underset{\sim}{x}'\underset{\sim}{d}_{\underset{\sim}{x}} = \underset{\sim}{x}'[\frac{1}{2}\underset{\sim}{b} + (B-\mu_o I_k)\underset{\sim}{x}] = \frac{1}{2}\underset{\sim}{x}'\underset{\sim}{b} + \underset{\sim}{x}'B\underset{\sim}{x} - \mu_o = 0$ holds among the elements of $\underset{\sim}{d}_{\underset{\sim}{x}}$ for any $\underset{\sim}{x}$ for which $\underset{\sim}{x}'\underset{\sim}{x} = 1$.

If we agree to make the standard assumptions in least squares theory concerning the independence and normality of errors, then it follows from what has been done so far that, under the hypotheses $H_{\underset{\sim}{x}_{FS}}$ and $H_{\underset{\sim}{x}_{SH}}$, $\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}$ and $\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$ will have singular multinormal distributions with $\underset{\sim}{0}$ mean vectors and dispersion matrices $V(\underset{\sim}{d}_{\underset{\sim}{x}_{FS}})$ and

$V(\underset{\sim}{d}_{\underset{\sim}{x}_{SH}})$, respectively. In particular, it can be shown fairly directly that the quantities

$$(2N/5\sigma^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{FS}}(I_2-\underset{\sim}{x}_{FS}\underset{\sim}{x}'_{FS})\underset{\sim}{d}_{\underset{\sim}{x}_{FS}} \text{ and } (2N/9\sigma^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{SH}}(I_3-\underset{\sim}{x}_{SH}\underset{\sim}{x}'_{SH})\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$$

are distributed as central χ^2 variates with one and two degrees of freedom, respectively. Of course, the idempotency of $(I_2-\underset{\sim}{x}_{FS}\underset{\sim}{x}'_{FS})$ and $(I_3-\underset{\sim}{x}_{SH}\underset{\sim}{x}'_{SH})$ is crucial to the argument.

Now, let s^2 be an estimate of σ^2 based on ν degrees of freedom which has been obtained either from a residual sum of squares (assuming an adequate model) or from a "pure error" sum of squares obtained by replication; moreover, we take s^2 to be distributed as a $\sigma^2\chi^2/\nu$ random variable independently of $\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}$ in the case of F. S. regression or of $\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$ in the case of S. H. regression. Then, it follows that $T_{\underset{\sim}{x}_{FS}} = (2N/5s^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{FS}}(I_2-\underset{\sim}{x}_{FS}\underset{\sim}{x}'_{FS})\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}$ is distributed as $F(1,\nu)$ and $T_{\underset{\sim}{x}_{SH}} = (N/9s^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{SH}}(I_3-\underset{\sim}{x}_{SH}\underset{\sim}{x}'_{SH})\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$ is distributed as $F(2,\nu)$, where $F(\gamma,\nu)$ denotes the F distribution with γ and ν degrees of freedom. If we now make use of the relations $\underset{\sim}{x}'_{FS}\underset{\sim}{d}_{\underset{\sim}{x}_{FS}} = 0$ and $\underset{\sim}{x}'_{SH}\underset{\sim}{d}_{\underset{\sim}{x}_{SH}} = 0$ to note that $T_{\underset{\sim}{x}_{FS}}$ can be written as $(2N/5s^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{FS}}\underset{\sim}{d}_{\underset{\sim}{x}_{FS}}$ and that $T_{\underset{\sim}{x}_{SH}}$ can be written as $(N/9s^2)\underset{\sim}{d}'_{\underset{\sim}{x}_{SH}}\underset{\sim}{d}_{\underset{\sim}{x}_{SH}}$, then it is reasonable to reject the hypotheses $H_{\underset{\sim}{x}_{FS}}: \underset{\sim}{\delta}_{\underset{\sim}{x}_{FS}} = \underset{\sim}{0}$ and $H_{\underset{\sim}{x}_{SH}}: \underset{\sim}{\delta}_{\underset{\sim}{x}_{SH}} = \underset{\sim}{0}$ at the α level of significance if $T_{\underset{\sim}{x}_{FS}} \geq F_\alpha(1,\nu)$ and $T_{\underset{\sim}{x}_{SH}} \geq F_\alpha(2,\nu)$, respectively, where $F_\alpha(\gamma,\nu)$ is the upper $100\alpha\%$ point of the $F(\gamma,\nu)$ distribution.

Finally, it follows from the structure of the above tests that the associated $100(1-\alpha)\%$ confidence regions for $\underset{\sim}{x}_o^*$ (corresponding to μ_o) for F. S. and S. H. regression have the following forms:

(11.1) $$R_{FS} = \{\underset{\sim}{x}_{FS}: \underset{\sim}{d}'_{\underset{\sim}{x}_{FS}}\underset{\sim}{d}_{\underset{\sim}{x}_{FS}} \leq 2.5s^2F_\alpha(1,\nu)/N\},$$

where $\underset{\sim}{d}_{\underset{\sim}{x}_{FS}} = \frac{1}{2}\underset{\sim}{b}_{FS}^{(1)} + (B_{FS}-\mu_o I_2)\underset{\sim}{x}_{FS}$; and,

$$\text{(11.2)} \qquad R_{SH} = \{\underset{\sim}{x}_{SH}: \underset{\sim}{d}'_{\underset{\sim}{x}_{SH}}\underset{\sim}{d}_{\underset{\sim}{x}_{SH}} \leq 9s^2F_\alpha(2,\nu)/N\},$$

where $\underset{\sim}{d}_{\underset{\sim}{x}_{SH}} = \frac{1}{2}\underset{\sim}{b}_{SH}^{(1)} + (B_{SH}-\mu_o I_3)\underset{\sim}{x}_{SH}$.

In what follows, we will attempt to provide a geometrical interpretation for the regions R_{FS} and R_{SH}. Now, if $\underset{\sim}{x}_o$ is a solution of (10.6) corresponding to a real root μ_o of (10.8), then $\underset{\sim}{d}_{\underset{\sim}{x}_o} = \underset{\sim}{0}$ by definition. So, the same procedures which led to (10.12) can also be applied here to obtain the expression

$$\text{(11.3)} \qquad \underset{\sim}{d}'_{\underset{\sim}{x}}\underset{\sim}{d}_{\underset{\sim}{x}} = \sum_{i=1}^{k}(\lambda_i-\mu_o)^2 w_i^2,$$

where the row vector $\underset{\sim}{w}' = (w_1, w_2,\ldots, w_k)$ is defined in (10.12).

Hence, from the form of (11.3), it follows that if $|C_o| = \Pi_{i=1}^{k}(\lambda_i-\mu_o)$ is non-zero, then $\underset{\sim}{d}'_{\underset{\sim}{x}}\underset{\sim}{d}_{\underset{\sim}{x}} = z$ represents the locus of an ellipsoid in the space of $x_1, x_2,\ldots, x_k$ for any fixed value of z, the center of this ellipsoid being at the stationary point $\underset{\sim}{x}_o$ and its principal axes lying directly on the set of coordinate axes $w_1, w_2, \ldots, w_k$.

In general, then, a point $\underset{\sim}{x}_{FS}$ will be in R_{FS} if and only if it lies not only on the unit circle but also on or within the ellipse $(\lambda_1-\mu_o)^2w_1^2 + (\lambda_2-\mu_o)^2w_2^2 = 2.5s^2F_\alpha(1,\nu)/N$, where $(w_1, w_2) = (\underset{\sim}{x}_{FS}-\underset{\sim}{x}_{FS}^{(o)})'P'_{FS}$, P_{FS} being the orthogonal matrix satisfying $P_{FS}B_{FS}P'_{FS} = \mathrm{diag}(\lambda_1, \lambda_2)$. Similarly, a point $\underset{\sim}{x}_{SH}$ will be in R_{SH} if and only if it lies not only on the unit sphere but also on or within the ellipsoid $(\lambda_1-\mu_o)^2w_1^2 + (\lambda_2-\mu_o)^2w_2^2 + (\lambda_3-\mu_o)^2w_3^2 = 9s^2F_\alpha(2,\nu)/N$, where $(w_1, w_2, w_3) = (\underset{\sim}{x}_{SH}-\underset{\sim}{x}_{SH}^{(o)})'P'_{SH}$, P_{SH} being the orthogonal matrix satisfying $P_{SH}B_{SH}P'_{SH} = \mathrm{diag}(\lambda_1, \lambda_2, \lambda_3)$.

Now, in (11.3), suppose that $(\lambda_1-\mu_o), (\lambda_2-\mu_o),\ldots, (\lambda_r-\mu_o)$, $r < k$, are small compared to the remaining $(k - r)$ characteristic roots of C_o. Then, the ellipsoid $\underset{\sim}{d}'_x\underset{\sim}{d}_x = z$ will be elongated in the direction of the axes $w_1, w_2,\ldots, w_r$. So, choices of $\underset{\sim}{x} = \underset{\sim}{x}_o + P'\underset{\sim}{w}$ such that $\underset{\sim}{x}'\underset{\sim}{x} = 1$ possibly differing markedly from the true value $\underset{\sim}{x}_o^*$ of the stationary point will give small values of $\underset{\sim}{d}'_x\underset{\sim}{d}_x$ over a wide range of values of the coordinates $w_1, w_2,\ldots, w_r$; such points would then most likely be included (and undesirably so) in confidence regions with structures like (11.1) and (11.2). In the limiting case when $(\lambda_1-\mu_o), (\lambda_2-\mu_o),\ldots, (\lambda_r-\mu_o)$ are exactly zero, C_o will be of rank $(k - r)$, and any point $\underset{\sim}{x} = \underset{\sim}{x}_o + P'\underset{\sim}{w}$ which satisfies (10.6) will do so independently of the values assigned to the coordinates $w_1, w_2,\ldots, w_r$.

An example of the construction of confidence regions of the type discussed in this section will be forthcoming.

12. Numerical Examples

Here, we will present two examples illustrating the use of the designs and techniques developed in the three previous sections. The first example is completely artificial.

Example I. Suppose that an S. H. model of order two has the form $\cos\phi\, P_{11}(\cos\theta) + \frac{1}{3}\cos2\phi\, P_{22}(\cos\theta) = \cos\phi\sin\theta + \cos2\phi\sin^2\theta$. To find the extremum points of this expression, we first note from (9.10) and (10.2) that $a_{00} = 0$, $\underset{\sim}{b}'_{SH} = (0, 1, 0)$, and $B_{SH} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$.

Using these results, it is easy to show that expression (10.10) takes the form $\mu^6 - \frac{9}{4}\mu^4 - \frac{1}{2}\mu^3 + \frac{3}{4}\mu^2 = 0$, which has double roots at 0 and -1 and single roots at 1/2 and 3/2. Since 3/2 is the largest root in this set and is obviously not a characteristic root of B_{SH} above, the theory previously developed tells us that the maximum value of $(\cos\phi \sin\theta + \cos 2\phi \sin^2\theta)$ over $0 \le \theta \le \pi$, $0 \le \phi \le 2\pi$, expressed in terms of the elements of x_{SH}, occurs at the unique point

$$\begin{pmatrix} \cos\theta \\ \cos\phi \sin\theta \\ \sin\phi \sin\theta \end{pmatrix} = \frac{1}{2}\begin{bmatrix} 3/2 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 5/2 \end{bmatrix}^{-1} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix},$$

or, in terms of θ and ϕ, at $\theta = \pi/2$ and $\phi = 0$. Since the smallest root in this set, namely -1, also turns out to be a characteristic root of B_{SH} above, the minimum of $(\cos\phi \sin\theta + \cos 2\phi \sin^2\theta)$, which is found by solving the pair of equations $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix} x_{SH} = -\frac{1}{2}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ and $x_{SH}'x_{SH} = 1$ simultaneously, occurs at the two distinct points $\frac{1}{4}(0, -1, \pm\sqrt{15})$, or, in terms of θ and ϕ, at $\theta = \pi/2$ and $\phi = \tan^{-1}(\pm\sqrt{15})$. The minimum is attained at more than one point due to the fact that $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ is singular.

This example illustrates the situation when a "singularity" arises (i.e., when a root of (10.8) is also a characteristic root of B), and, except for the purpose of crystallizing the general concepts presented in Section 10, is only of pedantic interest. The following example will be somewhat more realistic.

Example II. It was assumed that the true function representing a response surface was of the form $\eta(x_1,x_2;\beta) = 2 + 3x_1 - x_2 + 8x_1^2 + 5x_2^2 + 6x_1x_2$ and that the region of interest $\mathcal{X}$ was the circumference of a circle (scaled to unit radius). It was desired to graduate $\eta(x_1,x_2;\beta)$ by fitting by least squares an F. S. model of the appropriate degree. This fitting was to be accomplished by taking $r = 2$ observations at each of the $p = 6$ points $\phi_i = \pi i/3$, $i = 1, 2, \ldots, 6$. Note that this design is an F. S. optimal exact design of order two (see Theorem 9.2). Random normal deviates with $\sigma = 1$ were added to the values of $\eta(x_1,x_2;\beta)$ at the points of this design to generate the following set of observations: 13.620, 14.400 at $\phi = 0$; 10.552, 10.602 at $\phi = \pi/3$; 2.196, 3.696 at $\phi = 2\pi/3$; 6.390, 7.250 at $\phi = \pi$; 8.854, 10.684 at $\phi = 4\pi/3$; and, 5.408, 8.488 at $\phi = 5\pi/3$. The formulas (9.11)-(9.13) were then applied to this data to yield the estimates $a_0 = 8.512$, $a_1 = 3.198$, and $b_1 = -.922$. However, the sum of squares due to lack of fit of $\hat{y}_1(\phi;\underset{\sim}{b}_{FS})$, which was based on three degrees of freedom, was found to be significantly large when compared to the "pure error" sum of squares having six d.f. So, second-order terms were added in an attempt to provide a better graduation of the true response function and the resulting F. S. model of order two took the form $\hat{y}_2(\phi;\underset{\sim}{b}_{FS}) = 8.512 + 3.198\cos\phi - .922\sin\phi + 1.903\cos2\phi + 3.017\sin2\phi$. As can be seen from the ANOVA table below, there is no evidence of a lack of fit ($F(1,6) = 1.345$) of $\hat{y}_2(\phi;\underset{\sim}{b}_{FS})$.

SOURCE	D.F.	MEAN SQUARE
a_0	1	869.450
a_1	1	61.363
b_1	1	5.100
a_2	1	21.736
b_2	1	54.614
Lack of Fit	1	1.842
Pure Error	6	1.370
Total	12	

Confidence sets for those points on the circle at which $\eta(x_1,x_2;\beta)$ achieves its maximum and minimum are found as follows. For the values of a_1, b_1, a_2, b_2 above, the expression (10.9) evaluates to $\mu^4 - 28.21672\mu^2 - .02654\mu + 126.65651 = 0$, which has the four real roots -4.75491, -2.36725, 2.36529, and 4.75647. Since the two characteristic roots of $B_{FS} = \begin{bmatrix} 1.903 & 3.017 \\ 3.017 & -1.903 \end{bmatrix}$ are $\pm(a_2^2 + b_2^2)^{1/2} = \pm 3.567$, it follows that $\hat{y}_2(\phi;\underset{\sim}{b}_{FS})$ attains its maximum and minimum at the unique points

$$\underset{\sim}{x}_{min} = -\frac{1}{2}\left[\begin{pmatrix} 1.903 & 3.017 \\ 3.017 & -1.903 \end{pmatrix} + 4.75491\, I_2\right]^{-1}\begin{pmatrix} 3.198 \\ -.922 \end{pmatrix} = \begin{pmatrix} -.6020 \\ .7985 \end{pmatrix}, \text{ and}$$

$$\underset{\sim}{x}_{max} = -\frac{1}{2}\left[\begin{pmatrix} 1.903 & 3.017 \\ 3.017 & -1.903 \end{pmatrix} - 4.75647\, I_2\right]^{-1}\begin{pmatrix} 3.198 \\ -.922 \end{pmatrix} = \begin{pmatrix} .9351 \\ .3544 \end{pmatrix},$$

respectively; in terms of ϕ, then, the estimated locations of the points on the circle at which $\eta(x_1,x_2;\underset{\sim}{\beta})$ attains its maximum and minimum are at $\phi = .1153\pi$ and $.7944\pi$, respectively.

Now, it can be seen from the form of (11.1) that the $100(1-\alpha)\%$ confidence sets for the extremum points of $\eta(x_1,x_2;\underset{\sim}{\beta})$ can be specified exactly by determining the points of intersection of the ellipse $\underset{\sim}{d}'_{\underset{\sim}{x}_{FS}}\underset{\sim}{d}_{\underset{\sim}{x}_{FS}} = 2.5s^2F_\alpha(1,\nu)/N$ and the unit circle. These points of intersection are given by the real roots of the polynomial in $x_1 = \cos\phi$ of the general form

$$\begin{aligned}(12.1)\quad & 16\mu_o^2(a_2^2 + b_2^2)x_1^4 + 8\mu_o\{a_2[a_1(\mu_o - a_2) - b_1b_2] + b_2[b_1(\mu_o + a_2) \\ & - a_1b_2]\}x_1^3 + \{[a_1(\mu_o - a_2) - b_1b_2]^2 + [b_1(\mu_o + a_2) - a_1b_2]^2 \\ & - 8a_2\mu_o[(\mu_o + a_2)^2 + b_2^2 + \tfrac{1}{4}(a_1^2 + b_1^2) - k] - 16b_2^2\mu_o^2\}x_1^2 \\ & - 2\{[a_1(\mu_o - a_2) - b_1b_2][(\mu_o + a_2)^2 + b_2^2 + \tfrac{1}{4}(a_1^2 + b_1^2) - k] \\ & + 4b_2\mu_o[b_1(\mu_o + a_2) - a_1b_2]\}x_1 + [(\mu_o + a_2)^2 + b_2^2 \\ & + \tfrac{1}{4}(a_1^2 + b_1^2) - k]^2 - [b_1(\mu_o + a_2) - a_1b_2]^2 = 0,\end{aligned}$$

where $k = 2.5s^2F_\alpha(1,\nu)/N$. This polynomial will have only two real roots (except in the rare case when the ellipse is attenuated in such a way that it intersects the circle in more than two points), each root lying between -1 and +1.

Now, the value of k for a 95% confidence region in this example is $k = 2.5(1.370)(5.99)/12 = 1.71$. For this value of k and for the values of a_1, b_1, a_2, b_2 given earlier, expression (12.1) takes the form $4602.7558x_1^4 + 2145.4602x_1^3 - 1576.4737x_1^2 - 128.2284x_1 + 285.4421 = 0$ with roots $-.469212$, $-.722356$, and $.362721 \pm .226723i$ if $\mu_o = -4.75491$, and it takes the form $4605.7765x_1^4 - 950.3213x_1^3 - 6851.0484x_1^2 + 514.4296x_1 + 2722.0831 = 0$ with roots $.979843$, $.867891$, and $-.820701 \pm .146414i$ if $\mu_o = 4.75647$. Converting to radian measure, we find that the 95% confidence set for the maximum of $\eta(x_1,x_2;\underset{\sim}{\beta})$ on the circle is $\{\phi\colon .064\pi \leq \phi \leq .166\pi\}$ and for the minimum it is $\{\phi\colon .743\pi \leq \phi \leq .845\pi\}$. These sets do, in fact, include the true locations of the points on the circle at which $\eta(x_1,x_2;\underset{\sim}{\beta})$ achieves its maximum and minimum, namely $.126\pi$ and $.783\pi$, respectively.

We can slightly modify the above example to demonstrate the procedures that are to be used for the case of a response function defined on the real line. In particular, suppose that $\eta(x;\underset{\sim}{\beta}) = 2 + 3\cos\psi - \sin\psi + 8\cos^2\psi + 5\sin^2\psi + 6\sin\psi\cos\psi$, where $\psi = 2\pi(x-a)/(b-a)$, represents a response function which is to be graduated over the entire real line by an appropriately chosen F. S. model. We can take the region of interest to be the closed interval $[a,b]$ since $\eta(x;\underset{\sim}{\beta})$ is of (assumed to be known) period $(b-a)$. One can then show, if the random normal deviates are assigned as in the first part of this example, that exactly the same set of responses as before will be obtained, but, in this case, they will be associated with the six points $x_i = a + (b-a)\phi_i/2\pi = a + (b-a)i/6$, $i = 1, 2, \ldots, 6$, in the interval $[a,b]$. The fitted F. S. model of order two will have exactly the same form as $\hat{y}_2(\phi;\underset{\sim}{b}_{FS})$ above with ϕ replaced by ψ.

The 95% confidence intervals for the locations of the maximum and minimum of $\eta(x;\underset{\sim}{\beta})$ on $[a,b]$ will be

$\{x\colon a + .032(b-a) \leq x \leq a + .083(b-a)\}$, and

$\{x\colon a + .3715(b-a) \leq x \leq a + .4225(b-a)\}$, respectively.

The procedures developed for using S. H. models as graduating functions can be similarly applied to study response functions defined on the surface of a sphere or on the plane (see Section 1). The corresponding confidence regions for the stationary points of the response functions under study will, however, be somewhat more difficult to appreciate.

LIST OF REFERENCES

1. BOSE, R.C. and DRAPER, N.R. (1959). Second order rotatable designs in three dimensions. Ann. Math. Statist. 30 1097-1112.

2. BOX, G.E.P. and DRAPER, N.R. (1959). A basis for the selection of a response surface design. J. Amer. Statist. Assoc. 54 622-654.

3. BOX, G.E.P. and DRAPER, N.R. (1963). The choice of a second order rotatable design. Biometrika 50 335-352.

4. BOX, G.E.P. and HUNTER, J.S. (1954). A confidence region for the solution of a set of simultaneous equations with an application to experimental design. Biometrika 41 190-199.

5. BOX, G.E.P. and HUNTER, J.S. (1957). Multifactor experimental designs for exploring response surfaces. Ann. Math. Statist. 28 195-241.

6. BOX, G.E.P. and WILSON, K.B. (1951). On the experimental attainment of optimum conditions. J. Roy. Statist. Soc. Ser. B. 13 1-45.

7. DAVID, H.A. and ARENS, B.E. (1959). Optimal spacing in regression analysis. Ann. Math. Statist. 30 1072-1081.

8. DRAPER, N.R. and LAWRENCE, W.E. (1965). Designs which minimize model inadequacies: cuboidal regions of interest. Biometrika 52 111-118.

9. DRAPER, N.R. and LAWRENCE, W.E. (1965). Mixture designs for three factors. J. Roy. Statist. Soc. Ser. B. 27 450-465.

10. DRAPER, N.R. and LAWRENCE, W.E. (1965). Mixture designs for four factors. J. Roy. Statist. Soc. Ser. B. 27 473-478.

11. DRAPER, N.R. and LAWRENCE, W.E. (1967). Sequential designs for spherical weight functions. Technometrics 9 517-529.

12. ELFVING, G. (1959). Design of linear experiments. Probability and Statistics (Cramér Volume). Wiley, New York. 58-74.

13. FERGUSON, THOMAS S. (1967). Mathematical Statistics: A Decision Theoretic Approach. Academic Press, New York.

14. GARDINER, D.A., GRANDAGE, A.H.E., and HADER, R.J. (1959). Third order rotatable designs for exploring response surfaces. Ann. Math. Statist. **30** 1082-1096.

15. HADER, R.J., KARSON, M.J., and MANSON, A.R. (1969). Minimum bias estimation and experimental design for response surfaces. Technometrics **11** 461-475.

16. HANCOCK, HARRIS (1917). Theory of Maxima and Minima. Ginn and Company, Boston.

17. HOBSON, E.W. (1931). The Theory of Spherical and Ellipsoidal Harmonics. Cambridge University Press, London.

18. HOEL, P.G. (1965). Minimax designs in two-dimensional regression. Ann. Math. Statist. **36** 1097-1106.

19. KARLIN, SAMUEL and STUDDEN, WILLIAM, J. (1966). Optimal experimental designs. Ann. Math. Statist. **37** 783-815.

20. KIEFER, J. (1959). Optimal experimental designs. J. Roy. Statist. Soc. Ser. B. **21** 273-319.

21. KIEFER, J. (1960). Optimal experimental designs V, with applications to systematic and rotatable designs. Proc. Fourth Berkeley Symp. **1** 381-405.

22. KIEFER, J. and WOLFOWITZ, J. (1959). Optimum designs in regression problems. Ann. Math. Statist. **30** 271-294.

23. KIEFER, J. and WOLFOWITZ, J. (1960). The equivalence of two extremum problems. Canad. J. Math. **12** 363-366.

24. WALLACE, D. (1958). Intersection region confidence procedures with an application to the location of the maximum in quadratic regression. Ann. Math. Statist. **29** 455-475.