GEORGE ALEC RHODES TAYLOR

A dissertation on

EXPERIMENTAL DESIGNS AND ANALYSIS OF DATA FOR MIXTURES

Presented for the Degree of Master of Science
of the University of Aston in Birmingham, May 1972.

SUMMARY.

The problem of the design and analysis of experiments involving mixtures, where a restriction exists due to the fact that the sum of the proportions must be unity, was, as far as the writer is aware first mentioned in the literature by Claringbold*.

Scheffe seems to have been the first to go into the problem in any great detail and he obtained equations for analysing such experiments and also considered some of the ramifications such as process variables, fractionation and restrictions on composition which make it necessary to use pseudocomponents. The equations which Scheffe derived are examined and criticised and their use illustrated by examples from experiments in which the writer has been involved. Scheffe's method of dealing with process variables is also illustrated by a synthetic example.

Concentric shell designs recently developed by Doehlert are outlined and their relationship to balanced incomplete blocks, Scheffe's designs and Hadamard matrices is given.

Draper and Lawrence have developed more sophisticated designs using techniques worked out by Box and Draper for minimising variance and bias errors in regression equations. A critical comparison is made between their designs and Scheffe's.

Following Box's method of Evolutionary Operation, equations have been derived to enable the technique to be applied to production processes where mixtures are involved. This is illustrated by a simulated production process and some suggestions are made to overcome a difficulty encountered in this simulation which could easily occur in a real situation.

The Appendix contains original tables to enable confidence limits to be calculated when using Scheffe's designs for the special, but usual, case when the replicates are of equal size.

---

* Claringbold,P.J., "The Use of Simplex Designs in the Study of the Joint Action of Related Hormones", Biometrics, 1955, 11(2), 174.

TABLE OF CONTENTS.

TABLE OF CONTENTS (contd)

ACKNOWLEDGEMENTS.

It is a pleasure to acknowledge the help and advice given to me in the preparation of this Dissertation by several members of the staff of the Department of Mathematics of the University of Aston. Mr.W.Armstrong (now of the Production Engineering Department, of the University of Birmingham), because of his general interest in experimental design, was always ready to discuss any point that arose. Mr.I.G.Betteley clarified for me some of the more difficult parts of Professor H.Scheffe's papers. Mrs.S.M.Brookes similarly removed for me a stumbling block to my understanding of Professor N.Draper's and Dr.W.Lawrence's papers.

Without exception I have found authors of papers to whom I have written have been most helpful. I feel I must make special mention of Mr.J.W.Gorman of the Standard Oil Company and Professor N.Draper of the University of Wisconsin who have never failed to answer in detail my many queries.

Two members of the Mathematics Department of the University of Aston, without whose help I feel I could not have overcome the administrative problems associated with starting this Dissertation are Mr.N.Kerruish and Dr.A.M.Rajput.

On a different level, my wife's willingness for me to spend a great deal of time on this Dissertation instead of sharing the burden of "bringing up" our three daughters gave me the support which could not have been compensated for in any other way.

Introduction.

This dissertation examines and illustrates the methods which have been developed by Scheffe (1958,1963) and others for the quantitative examination of mixtures. Most of the information is contained in two papers by Scheffe, two by Lambrakis (1968a,1968b) and one by Gorman and Hinman (1962). The two papers by Scheffe and Lambrakis are very erudite and mathematically elegant. But they can probably be read and understood only by the mathematically sophisticated and those to whom the exercise is a satisfying end in itself. Many of those who are in a position to put the techniques to use at the "shop floor level" are probably the ones who have the least time or detailed mathematical knowledge necessary to bring them down to practical terms. Gorman and Hinman have gone some way to remedy this in their paper in which they illustrate with practical examples the main contents of Scheffe's first paper.

In this dissertation the writer hopes to amplify the work done by Gorman and Hinman and to take it several stages further illustrating some of the more advanced techniques by means of examples.

The Problem.

There are many situations where the response does not depend on the total amount of a mixture present (an extensive property) as would be the case in a field experiment with fertilisers, but only on the proportions of the components present (an intensive property) as would be the case in the octane rating of a blend of petrols or in an animal feeding experiment where certain dietary combinations had to be tried out. In the latter case if the usual analysis of variance techniques were applied one would find oneself in the position

of trying to get an animal to eat twice or three times as much as the others. Even if this could be done the results would be of doubtful value.

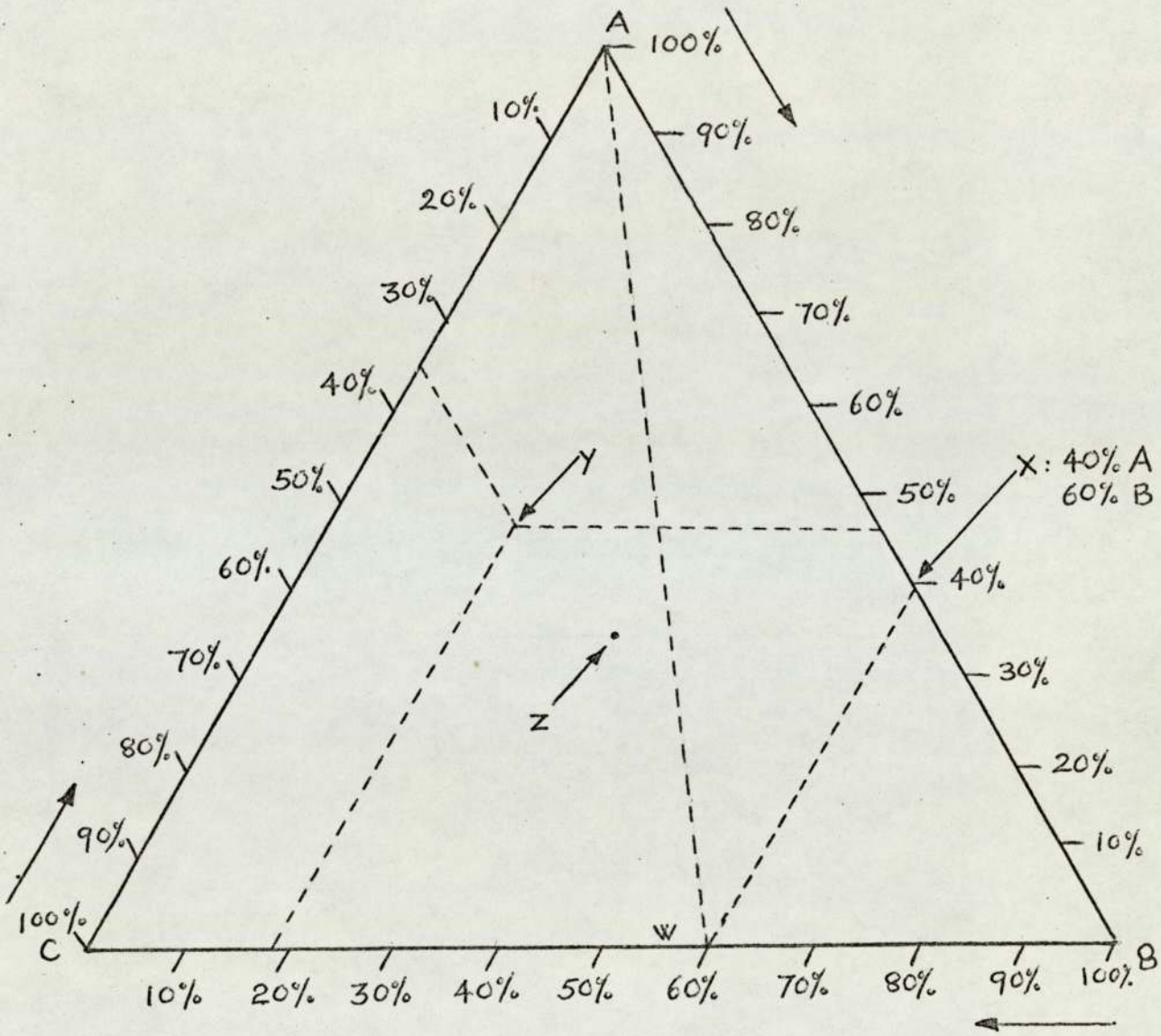
Similarly, if one wished to investigate the tensile strength of, say, stainless steel which has an approximate analysis of 18% chromium, 8% nickel and 74% iron and it was decided in the first instance to increase the chromium content by 2%; does one reduce the nickel by 2% or the iron by 2% or reduce both by 1%? Are the observed changes in tensile strength due to an increase in chromium or to a decrease in the iron or the nickel or to both? It can be appreciated that the usual factorial analysis no longer applies. Contrast this situation with a similar experiment which one might carry out on a steel whose phosphorus and sulphur content are of the order of 0.001%. Any change in the latter would bring about a negligible change in the overall iron content. Similarly, the application of fertilisers at a few ounces per square yard does not produce any appreciable change in the soil concentration. Scheffe's method recognises that for systems involving composition the sums of the proportions by weight, volume etc., must sum to unity. Therefore, the factor space is a regular simplex. For three components this will be an equilateral triangle; for five components a regular tetrahedron; for four or more components no geometric model is possible.

Simplex Lattices.

Any combination of three materials can be represented by a point in an equilateral triangle, a concept used extensively by chemists and metallurgists in plotting phase diagrams. The method is illustrated in Fig.1. A point on

A TWO DIMENSIONAL SIMPLEX LATTICE

Fig 1



the apex A represents 100% of component A. A point on the line A-B represents a binary miexture of A and B and none of C. A point within the triangle consists of a ternary mixture of A, B and C. Thus, the point X represents 40% of A and 60% of B; the point Y represents 46% of A, 17% of B and 37% of C; the point Z represents equal amounts of A, B and C and is the centroid of the triangle. The idea can be extended to four components in which case a quaternary mixture would be represented by a point within the tetrahedron. Notice that any point on the line A-W represents a constant ratio of B to C in the case illustrated this will be B:C::6:4.

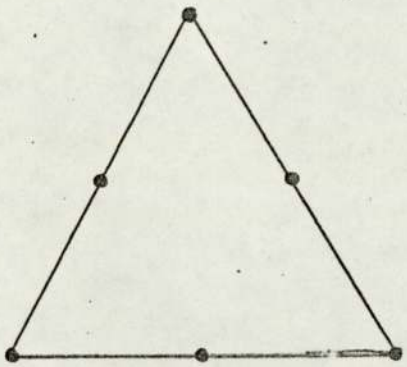
A (q,m) simplex lattice is defined as a lattice involving q components to which a polynomial of degree m is to be fitted. Examples of different lattices for q=3 are given in Fig. 2. The special cubic lattice is made by adding a centre point to the two dimensional face of the quadratic lattice. It will be shown later that this simplex, called a simplex centroid design, has several desirable properties which make its use preferable in many cases to the other designs shown in Fig. 2.

The proportions of the components are 0, 1/m,2/m,...,1 giving a total of m+1 equally spaces values from 0 to 1 and all possible mixtures with these proportions are used. For the quadratic lattice (q,2) the proportions are $0,\frac{1}{2},1$ and mixtures are all the possible permutations of (0,0,1) and of $(0,\frac{1}{2},\frac{1}{2})$. For the cubic lattice (q,3) the proportions are $0,\frac{1}{3},\frac{2}{3},1$ and the mixtures are all the permutations of (0,0,1), all the permutations of $(0,\frac{1}{3},\frac{2}{3})$ and the centre point $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$

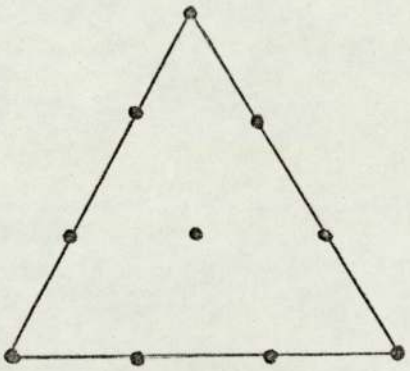
Table 1 gives the number of points required for any lattice (except the special cubic) and is given by Scheffe (1958) as $(m+q-1)!/m!(q-1)!$ For the special cubic the formula is

EXAMPLES OF LATTICE DESIGNS.

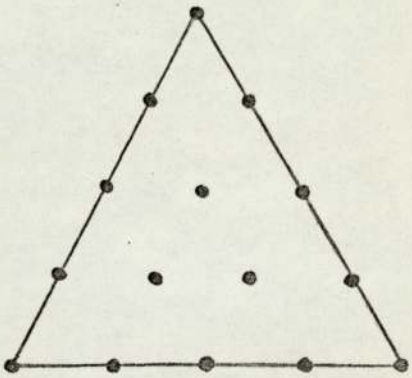
Fig 2



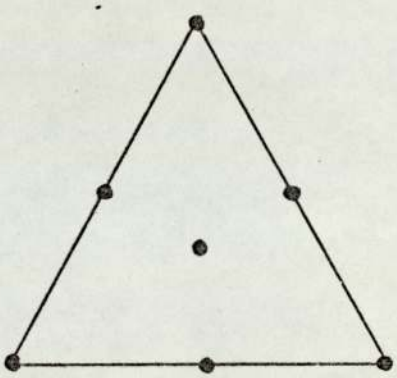
QUADRATIC (3,2)



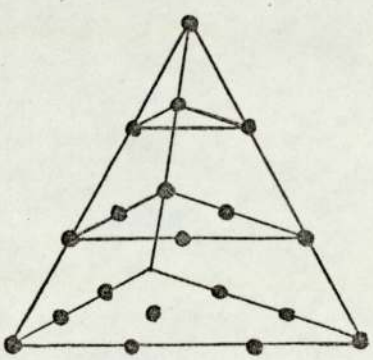
CUBIC (3,3)



QUARTIC (3,4)



SPECIAL CUBIC



CUBIC (4,3)

$\{3q(q+1) + q(q-1)(q-2)\}/6$. Doehlert (1970) has worked out alternative formulae in the form of nested sums and has given an algorithm for calculating the required figures. Doehlert's formulae are,

$$m = 2 \qquad q(q+1)/2$$

$$m = 3 \qquad \sum_{i=1}^{q} i(i+1)/2$$

$$m = 4 \qquad \sum_{j=1}^{q} \sum_{i=1}^{i} i(i+1)/2 \qquad \text{etc.}$$

Table 1 gives the number of points for various values of q and m. Doehlert's* algorithm is that to obtain any figure one adds the one above to the one on the left. Thus 2002 in row 8 and column 4 is obtained by 1287 + 715 = 2002. This does not hold, of course, for the case of the special cubic.

TABLE 1.

Number of Mixtures in Various Types of Lattices

No. of components q	m: 2	3	4	5	Special Cubic
3	6	10	15	21	7
4	10	20	35	56	14
5	15	35	70	126	25
6	21	56	126	252	41
7	28	84	210	462	63
8	36	120	330	792	92
9	45	165	495	1287	129
10	55	220	715	2002	175

It is surprising that Scheffe does not mention it and it is obvious from Doehlert's* paper that he (Doehlert) has not appreciated it, that the figures in Table 1 are merely those

*Private communication.

in Pascal's triangle with the first two and last two figures struck out! Further, when m=2, the numbers generated are the well-known triangular numbers for which Eperson (1967) has deduced several interesting relationships.

Relation of Lattice Designs to Uniform Shell Designs.

Doehlert (1967) has developed a new series of experimental designs which have an equally spaced distribution of points lying on concentric spherical shells. These have uniform space filling properties and show an interesting relationship with Scheffe's lattice designs and with well-known unsolved problems in balanced incomplete block designs and Hadamard matrices (C.L.Lui.) For instance, Hadamard* matrices of order n are known to exist when n: 1,2,4, ..., 4j ($j \leq 50$) with the exception j = 47 i.e. n = 188. If such a matrix is normalised by arranging the first row and first column to consist of +1's, and if this row and column are struck out and the -1's replaced by zeros we have a symmetrical balanced incomplete block. However, Hadamard matrices for $j > 50$ can only be conjectured. Consequently, symmetrically balanced incomplete block designs for these values are similarly open to conjecture.

A regular simplex in 2-space can be defined in terms of three cartesian coordinates A, B and C

A (0.000, 0.000)

B (1.000, 0.000)

C (0.500, 0.866)

These points are labelled in Fig. 3

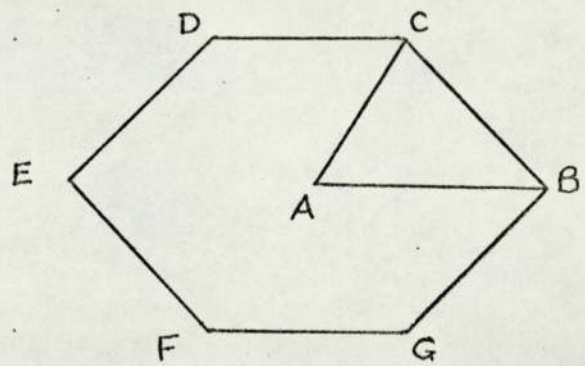


Fig 3

*See Appendix 1

If each point is subtracted from each other point four more points are obtained which define a regular hexagon

A-B	(-1.000, 0.000)	E
A-C	(-0.500, -0.866)	F
C-B	(-0.500, 0.866)	D
B-C	(0.500, -0.866)	G

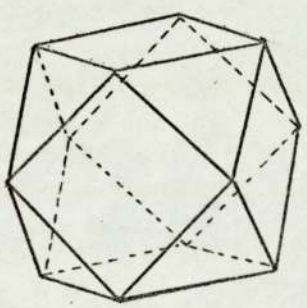
For three factors the starting point is a 3-space simplex i.e. a regular tetrahedron

A	(0.000,	0.000,	0.000)
B	(1.000,	0.000,	0.000)
C	(0.500,	0.866,	0.000)
D	(0.500,	0.289,	0.816)

If these points are subtracted from each other we obtain:

A-B	(-1.000,	0.000,	0.000)
A-C	(-0.500,	-0.866,	0.000)
A-D	(-0.500,	-0.289,	-0.816)
B-C	(0.500,	-0.866,	0.000)
B-D	(0.500,	-0.866,	-0.816)
C-B	(-0.500,	0.866,	0.000)
C-D	(0.000,	0.577,	-0.816)
D-B	(-0.500,	0.289,	0.816)
D-C	(0.000,	-0.577,	0.816)

These nine points together with the four original points define a cuboctahedron with a centre point



Cuboctahedron

For d factors a regular simplex is formed by adding to the d - 1 simplex the point

$$\frac{1}{2}, \frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{6}}, \ldots\ldots \frac{1}{\sqrt{\{2(d-1)(d-2)\}}}, \frac{1}{\sqrt{\{2d(d-1)\}}}, \frac{\sqrt{(d+1)}}{\sqrt{(2d)}}$$

To use the formula one works from right to left noticing that only the last two coordinates change and that the last coordinate is always $\sqrt{(d+1)}/\sqrt{(2d)}$.

Example.

To find the basic coordinates when (a) d = 4 and (b) d = 5.

a) d = 4 $\quad \frac{(4+1)}{(2\times4)} = 0.791$ and $\frac{1}{2\sqrt{6}} = 0.204$ etc.

giving (0.500, 0.289, 0.204, 0.791)

b) d = 5 $\quad \frac{\sqrt{(5+1)}}{\sqrt{(2\times5)}} = 0.775, \quad \frac{1}{\sqrt{\{2\times5(5-1)\}}} = 0.158$ etc.

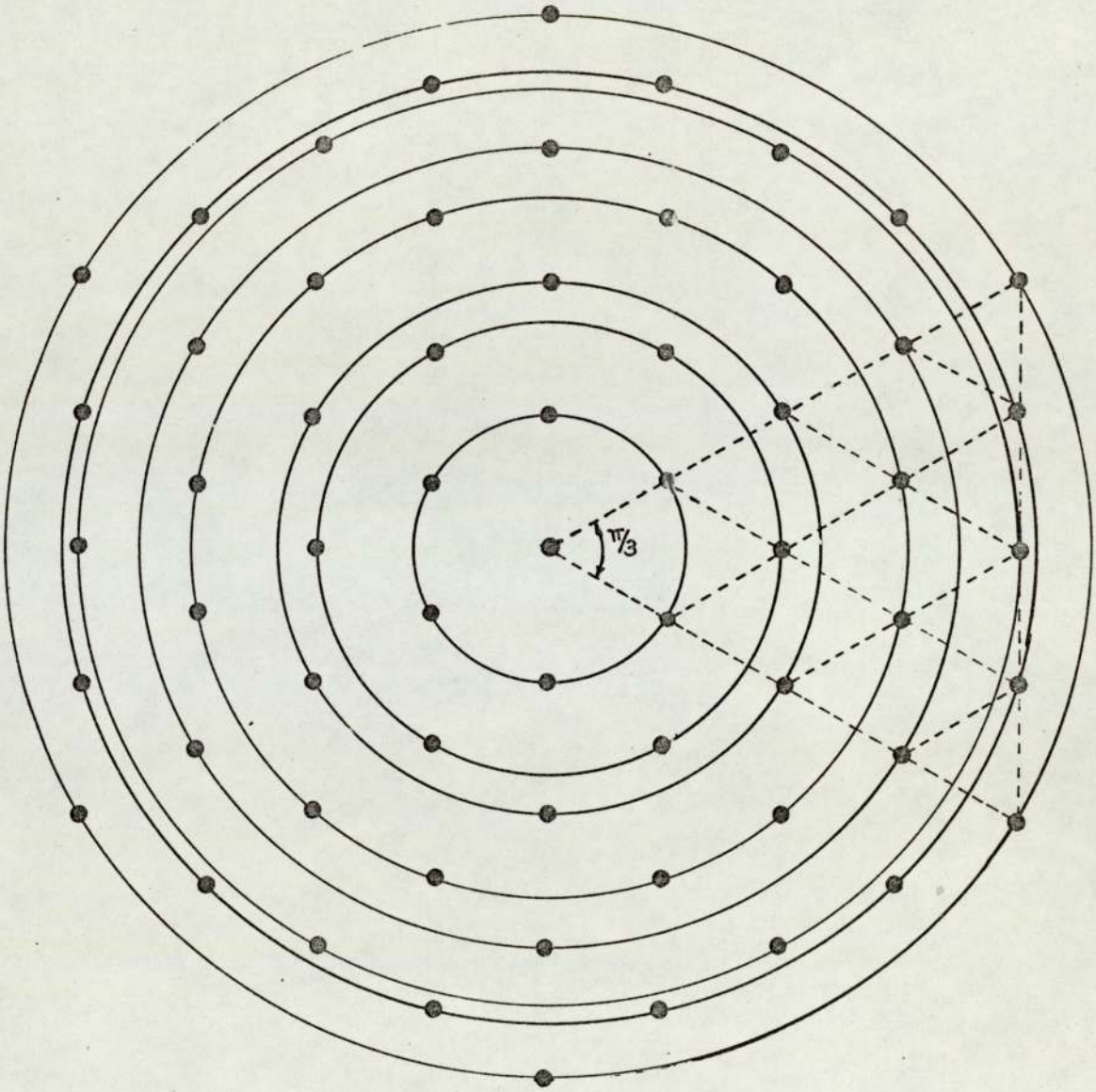
giving (0.500, 0.289, 0.204, 0.158, 0.775)

Putting these results in the form of a table together with those previously obtained we have

0.000	0.000	0.000	0.000	0.000
1.000	0.000	0.000	0.000	0.000
0.500	0.866	0.000	0.000	0.000
0.500	0.289	0.816	0.000	0.000
0.500	0.289	0.204	0.791	0.000
0.500	0.289	0.204	0.158	0.775

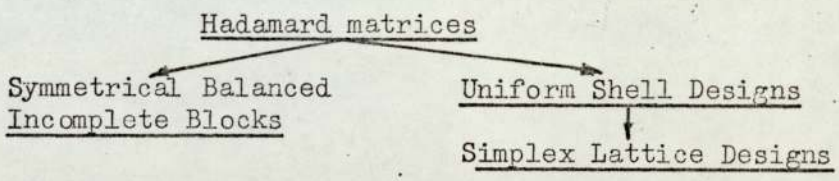
From these figures the coordinates of uniform shell designs up to d = 5 may be calculated. In Doehlert's paper (1970) more detailed figures up to d = 10 are tabulated. However, Doehlert and Klee (1970) have shown that the levels of these designs can generally be reduced by rotating the design and have calculated tables giving designs up to and including 14 factors.

Reverting to Doehlert's designs and their



First Eight Shells of the Rhombic Lattice.
Fig 4.

generation: if the same quantities subtracted in the process outlined above are now both subtracted from and added to the points in the uniform shell design, larger designs are obtained. This process carried out to a radius of 4 in 2-space produces the design in Fig. 4. The dots are at the intersections of equally spaced lines, at 60° to each other i.e. a rhombic or $\pi/3$ lattice. The points which have been joined by broken lines are the ones of interest. Working from the centre-point outwards it will be found that the developing pattern generates sequentially the simplexes necessary for the development of linear, quadratic, cubic, quartic etc., response surfaces (see Fig. 2) thus showing an interesting relationship between two apparently unconnected experimental designs. If one goes back a step further one can regard Hadamard matrices as a common souce of three types of experimental design.



Reverting now to the simplex designs; at some stage a decision has to be made what the value of m shall be in a given experiment. A decision will be reached base on

a) the maximum number of experiments which can be accommodated bearing in mind the need for replication if a measure of error is to be obtained

b) the adequacy of the polynomial chosen. A quadratic polynomial describes a response surface with no more than one maximum or one minimum, but not both, and with no point of inflexion; a cubic polynomial will give a maximum and a minimum or a point of inflexion and so on.

Bearing these factors in mind the writer feels

that a cubic polynomial is sufficiently accurate for most practical purposes while not involving a prohibitively large number of experiments.

The foregoing remarks will have to be modified when fractionation of simplex centroid designs are discussed which allow one to reduce the number of experimental points with the accompanying restrictions which fractionation involves.

Polynomials on the Simplex.

We will show how some of the equations used in this technique are derived.

A polynomial of degree n in q variables $x_1, x_2, \ldots, x_q$ subject to the restriction

$$x_1 + x_2 + \ldots + x_q = 1 \qquad (1)$$

will be of the form

$$y = b_0 + \sum_{1 \leq i \leq q} b_i x_i + \sum_{1 \leq i \leq j \leq q} b_{ij} x_i x_j + \sum_{1 \leq i \leq j \leq k \leq q} b_{ijk} x_i x_j x_k + \ldots$$

Consider, in the first case, where there are 3 components i.e. $q = 3$ and we are interested in a second order polynomial

$$y = b_0 + b_1 x_1 + b_3 x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2 \qquad (2)$$

Now, by (1), $x_1 + x_2 + x_3 = 1$

and $$b_0 = b_0 x_1 + b_0 x_2 + b_0 x_3 \qquad (3)$$

Hence b_0 may be eliminated from (2) by substituting (3)

$$y = (b_0+b_1)x_1 + (b_0+b_2)x_2 + (b_0+b_3)x_3 + b_{12} x_1 x_2 + b_{13} x_1 x_3 + b_{23} x_2 x_3 + b_{11} x_1^2 + b_{22} x_2^2 + b_{33} x_3^2$$

The squared terms may be eliminated by multiplying (3) by x_1, x_2 and x_3 to form the identities

$$\left.\begin{aligned} x_1^2 &= x_1 - x_1x_2 - x_1x_3 \\ x_2^2 &= x_2 - x_1x_2 - x_2x_3 \\ x_3^2 &= x_3 - x_1x_3 - x_2x_3 \end{aligned}\right\}$$

Substituting we obtain

$$y = (b_o+b_1+b_{11})x_1+ (b_o+ b_2+ b_{22})x_2+(b_o+b_2+b_{33})x_3$$
$$+(b_{12}-b_{11}-b_{22})x_1x_2+(b_{13}-b_{11}-b_{33})+(b_{23}-b_{22}-b_{33})x_2x_3$$

Replacing the sums of the constants by

$$\beta_1 = b_o + b_1 + b_{11} \quad \text{etc.},$$

yields

$$y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{22}x_2x_3$$

or, more generally

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j$$

If a third order polynomial is required nine more terms must be added to give

$$\begin{aligned} y = b_o + \ldots\ldots &+ b_{122}x_1x_2^2 + b_{112}x_1^2x_2 \\ &+ b_{133}x_1x_3^2 + b_{113}x_1^2x_3 \\ &+ b_{233}x_2x_3^2 + b_{223}x_2^2x_3 \\ &+ b_{123}x_1x_2x_3 \\ &+ b_{111}x_1^3 + b_{222}x_2^3 + b_{333}x_3^3 \end{aligned}$$

Proceeding as before and using the following three equations derived from above to eliminate x^3

$$\left.\begin{aligned} x_1^3 &= x_1^2 - x_1^2x_2 - x_1^2x_3 = x_1-x_1x_2-x_1x_3-x_1^2x_2-x_1^2x_3 \\ x_2^3 &= x_2^2 - x_1x_2^2 - x_2^2x_3 = x_2-x_1x_2-x_2x_3-x_1x_2^2-x_2^2x_3 \\ x_3^3 &= x_3^2 - x_2x_3^2 - x_2x_3^2 = x_3-x_1x_3-x_2x_3-x_1x_3^2-x_2x_3^2 \end{aligned}\right\}$$

we obtain

$$\begin{aligned} y = \beta_1x_1 &+ \beta_2x_2 + \beta_3x_3 \\ &+ \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_2 \\ &+ \beta_{112}x_1^2x_2 + \beta_{113}x_1^2x_3 + \beta_{223}x_2^2x_3 \\ &+ \beta_{122}x_1x_2^2 + \beta_{133}x_1x_3^2 + \beta_{233}x_2x_3^2 \\ &+ \beta_{123}x_1x_2x_3 \end{aligned} \qquad (4)$$

Now, from considerations of symmetry, $\beta_{iij} = -\beta_{ijj} = \gamma_{ij}$, say, which enables us to write terms involving coefficients of this type in a more compact form

$$\begin{aligned}&\beta_{iij}x_i^2x_j + \beta_{ijj}x_ix_j^2\\ &= \beta_{iij}x_i^2x_j - \beta_{iij}x_ix_j^2\\ &= \gamma_{ij}x_ix_j(x_i - x_j) \qquad (j > i)\end{aligned}$$

Hence substituting in (4)

$$\begin{aligned}y = {} & \beta_1x_1 + \beta_2x_2 + \beta_3x_3\\ & + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3\\ & + \gamma_{12}x_1x_2(x_1-x_2)+\gamma_{13}x_1x_3(x_1-x_3)+\gamma_{23}x_2x_3(x_2-x_3)\\ & + \beta_{123}x_1x_2x_3 \qquad (5)\end{aligned}$$

This third order equation may be written more generally as

$$y = \sum_{1\leqslant i\leqslant q}\beta_ix_i + \sum_{1\leqslant i<j\leqslant q}\beta_{ij}x_ix_j + \sum_{1\leqslant i<j\leqslant q}\gamma_{ij}x_ix_j(x_i-x_j)$$

$$+ \sum_{1\leqslant i<j<k\leqslant q}\beta_{ijk}x_ix_jx_k \qquad (6)$$

The equation for the special cubic (simplex centroid) is

$$y = \sum_{1\leqslant i\leqslant q}\beta_ix_i + \sum_{1\leqslant i<j\leqslant q}\beta_{ij}x_ix_j + \sum_{1\leqslant i<j<k\leqslant q}\beta_{ijk}x_ix_jx_k$$

Calculation of the coefficients of the polynomials.

The coefficients for the cubic equation (5) will be derived and those for the quadratic and special cubic quoted.

Let the response of the pure component be y_i. Putting $x_i = 1$ (and hence $x_j = x_k = 0$) and inserting this value in (6) we find

$$\beta_i = y_i$$

i.e. the coefficient associated with the pure components is the value of the response obtained with the pure component, a result which is generally true for all the equations of whatever degree.

Let the response of the binary mixtures be y_{iij} where the significance to be attached to the subscript is that at that point the mixture consists of 2 parts of component i and 1 part component j. At the point $x_i = \frac{2}{3}$, $x_j = \frac{1}{3}$ from (6)

$$y_{iij} = \tfrac{2}{3}\beta_i + \tfrac{1}{3}\beta_j + \tfrac{2}{9}\beta_{ij} + \tfrac{2}{27}\gamma_{ij} \qquad (7)$$

and at the point $x_i = \frac{1}{3}$, $x_j = \frac{2}{3}$

$$y_{ijj} = \tfrac{1}{3}\beta_i + \tfrac{2}{3}\beta_j + \tfrac{2}{9}\beta_{ij} + \tfrac{2}{27}\gamma_{ij} \qquad (8)$$

Adding (7) and (8)

$$y_{ijj} + y_{iij} = \beta_i + \beta_j + \tfrac{4}{9}\beta_{ij}$$

Substituting and rearranging

$$\beta_{ij} = \tfrac{9}{4}(y_{iij} + y_{ijj} - y_i - y_j)$$

Subtracting (7) and (8)

$$y_{ijj} - y_{ijj} = \tfrac{1}{3}\beta_j - \tfrac{1}{3}\beta_i + \tfrac{4}{27}\gamma_{ij}$$

Substituting and rearranging

$$\gamma_{ij} = \tfrac{9}{4}(3y_{iij} - 3y_{ijj} - y_i + y_j)$$

Finally, to evaluate β_{ijk}, put $x_i = x_j = x_k = \frac{1}{3}$ in (6)

$$y_{ijk} = \tfrac{1}{3}\beta_i + \tfrac{1}{3}\beta_j + \tfrac{1}{3}\beta_k + \tfrac{1}{9}\beta_{ij} + \tfrac{1}{9}\beta_{ik} + \tfrac{1}{9}\beta_{jk} + \tfrac{1}{27}\beta_{ijk}$$

(All γ terms are zero at this point)

Using the above results we obtain

$$\beta_{ijk} = 27y_{ijk} - (27/4)(y_{iij} + y_{ijj} + y_{iik} + y_{ikk} + y_{jjk} + y_{jkk}) + (9/2)(y_i + y_j + y_k)$$

The corresponding formulae for the quadratic,

special cubic and higher order polynomials may be similarly computed. Those for the first three models are quoted below:-

Quadratic Model.

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j$$

$$\beta_i = y_i$$
$$\beta_{ij} = 4y_{ij} - 2y_i - 2y_j \qquad (9)$$

Special Cubic Model.

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j + \sum_{1\leq i<j<k\leq q} \beta_{ijk} x_i x_j x_k$$

$$\beta_i = y_i$$

$$\beta_{ij} = 4y_{ij} - 2y_i - 2y_j$$

$$\beta_{ijk} = 27y_{ijk} - 12(y_{ij}+y_{ik}+y_{jk})+3(y_i+y_j+y_k)$$

Cubic Model.

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j + \sum_{1\leq i<j\leq q} \gamma_{ij} x_i x_j (x_i - x_j) + \sum_{1\leq i<j<k\leq q} \beta_{ijk} x_i x_j x_k$$

$$\begin{aligned}
\beta_i &= y_i \\
\beta_{ij} &= (9/4)(y_{iij}+y_{ijj}-y_i-y_j) \\
\gamma_{ij} &= (9/4)(3y_{iij}-3y_{ijj}-y_i+y_j) \\
\beta_{ijk} &= 27y_{ijk}-(27/4)(y_{iij}+y_{ijj}+y_{iik}+y_{ikk}+y_{jjk}+y_{jkk}) \\
&\qquad +(9/2)(y_i+y_j+y_k)
\end{aligned} \qquad (10)$$

Variance of the Predicted Response.

The method of deriving the formula necessary for the calculation of the variance will be given for the cubic equation and those for the quadratic and special cubic quoted.

We replace the β's and γ's by their estimates to give the predicted response $\tilde{y}$.

$$\tilde{y} = y_1x_1 + y_2x_2 + y_3x_3$$
$$+ (9/4)(y_{112} + y_{122} - y_1 - y_2)x_1x_2 + \ldots.$$
$$+ (27y_{123} - (27/4)(y_{112}+y_{122}+y_{113}+y_{133}+y_{223}+y_{233})$$
$$+ (9/2)(y_1+y_2+y_3)x_1x_2x_3$$

We now separate out the various coefficients

1) Coefficient of y_1 , C_1 say.

$$C_1 = x_1 - (9/4)x_1x_2 - (9/4)x_1x_3 - (9/4)x_1x_2(x_1-x_2)$$
$$- (9/4)x_1x_3(x_1-x_3) + (9/4)x_1x_2x_3$$

Factorising and using $1-x_3 = x_1+x_2$

$C_1 = (x_1/2)(3x_1-1)(3x_1-2)$ and, in general

$$C_i = (x_i/2)(3x_i-1)(3x_i-2).$$

2) Coefficient of y_{112}, C_{112}, say

$$C_{112} = (9/4)x_1x_2 + (27/4)x_1x_2(x_1-x_2) - (27/4)x_1x_2x_3$$
$$= (9/2)x_1x_2(3x_1-1)$$ and, in general, $C_{iij} = (9/2)x_ix_j(3x_i-1)$

3) Coefficient of y_{122}, say

$C_{122}(9/2)x_1x_2(3x_2-1)$ and, in general, $C_{ijj} = (9/2)x_ix_j(3x_j-1)$

4) Coefficient of y_{123}, C_{123}, say

This is easily seen to be $C_{123} = 27x_1x_2x_3$

and, in general $C_{ijk} = 27x_ix_jx_k$

Hence, for the cubic polynomial

$$\tilde{y} = \sum_{1\leq i\leq q} C_i y_i + \sum_{1\leq i<j\leq q} (C_{iij} y_{iij} + C_{ijj} y_{ijj}) + \sum_{1\leq i<j<k\leq q} C_{ijk} y_{ijk}$$

Using the form of the equation given on Page 14 where C_i, c_{iij}, C_{ijj} and C_{ijk} have the values given above, we can determine the variance of the predicted response. If the variance of the simplex is σ^2 then the variance of the predicted response is given by

$$\mathrm{Var}(\tilde{y}) = \sigma^2 \left\{ \sum_{1\leq i\leq q} \frac{C_i^2}{r_i} + \sum_{1\leq i<j\leq q} \frac{C_{iij}^2}{r_{iij}} + \sum_{1\leq i<j\leq q} \frac{C_{ijj}^2}{r_{ijj}} + \sum_{1\leq i<j<k\leq q} \frac{C_{ijk}^2}{r_{ijk}} \right\}$$

where r_i, r_{iij}, r_{ijj} and r_{ijk} are the number of observations on y_i, y_{ij} and y_{ijk}.

Each of the polynomials are orthogonal on the lattice in the sense that each equals unity at the lattice point associated with it. For example, for pure components when $x_1=1$ we have $(\frac{1}{2})(3-1)(3-2) = 1$; for binary mixtures when $x_1 = \frac{2}{3}$ and $x_2 = \frac{1}{3}$ we have $(\frac{2}{2})(\frac{2}{3})(\frac{1}{3})(3x\frac{2}{3}-1) = 1$ and for ternary mixtures when $x_1 = x_2 = x_3 = \frac{1}{3}$ we have $27(\frac{1}{3})(\frac{1}{3})(\frac{1}{3}) = 1$.

Allocation of measurements on the lattice.

It is obviously highly desirable to know how the experimental observations should be distributed over the lattice to give the minimum variance. A foreknowledge of this would enable an experimenter to allocate his, perhaps limited, resources to the best advantage.

This can be done by taking the number of observations at each lattice point proportional to the maximum of the squares of the coefficients. This allocates to each observed mean a number of observations such that the

maximum contribution to the variance of the predicted response is the same at each point. We have, therefore, to calculate the maximum values of the C coefficients.

1) Maximum value of C_i^2

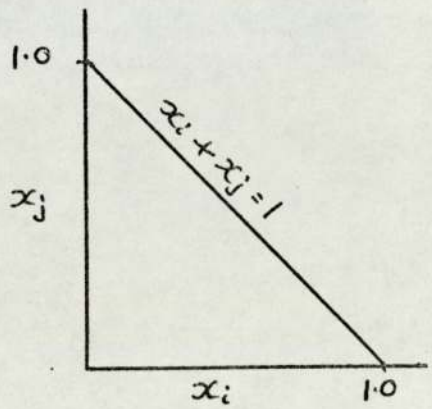
By rearranging C_i given on page 18

$$C_i = (9/2)(x_i^3 - x_i^2) + x_i$$

The maximum value of this is obviously $x_i = 1$ and the minimum at $x_i = 0$. At the maximum value of x_i, $C_i = 1$ and hence $\max(C_i^2) = 1$.

2) Maximum value of C_{ijj}^2

C_{ijj} is a function of x_i and x_j only and the values of x_i and x_j must lie in the triangle $x_i \geq 0$, $x_j \geq 0$, $x_i + x_j \leq 1$.



Now,

$C_{ijj}^2 = (81/4)x_i^2 x_j^2 (3x_i - 1)^2$

and the maximum value is when $x_i + x_j = 1$.

Under this restriction

$C_{ijj} = (9/2)x_i(1-x_i)(3x_i-1)$

Differentiation shows that the maximum is at $x_i = (4-\sqrt{7})/9$ corresponding to $C_{ijj} = (10+7\sqrt{7})/27$ and hence to $\max(C_{iij}^2) = (443+140\sqrt{7})/729 = 1.116$. The same result holds for C_{iij}^2.

3) The maximum value of C_{ijk}^2

This is obviously when $x_i = x_j = x_i = \frac{1}{3}$ at which point $\max(C_{ijk}^2) = 1$.

3) contd.

Therefore, the observations must be allocated amongst the points in the proportion

$$C_i : C_{ijj} : C_{ijk} = 1 : 1.116 : 1$$

This, for all practical purposes, is to say that one must take equal numbers of observations at each lattice point, a result which also holds for the quadratic response surface.

Allocation of measurements on the lattice for the case of the special cubic.

The way in which one tests the adequacy of fit of the polynomials is by taking extra observations (check points) at points of particular interest and comparing these observed values with those calculated from the polynomial.

Referring to Fig. 2 ,it would be reasonable if one were fitting a second order polynomial, to take one check point at the centre of the lattice. Then, if the second order polynomial was found not to give a good fit when measured against this centre check point, the check point itself could be incorporated in the equation to give a third order polynomial which would be a better fit. This technique is known as "augmentation" and the special cubic lattice is sometimes referred to as an "augmented" cubic lattice.

If one decides to make observations at the seven lattice points of the special cubic to fit a third order polynomial, then the remarks of the previous section apply; equal numbers of observations at all points. However, if used as a check point to test the goodness of fit of a second order polynomial the distribution of observations over the lattice works out differently.

Suppose we have derived a quadratic polynomial; at the centre point $x_i = x_j = x_k = \frac{1}{3}$ the response is given by

$$y_{ijk} = \tfrac{1}{3}\beta_i + \tfrac{1}{3}\beta_j + \tfrac{1}{3}\beta_k + \tfrac{1}{9}\beta_{ij} + \tfrac{1}{9}\beta_{ik} + \tfrac{1}{9}\beta_{jk}$$

Now, for the quadratic polynomial the coefficients in terms of the observed responses are given by

$$\left.\begin{aligned} \beta_i &= y_i \\ \text{and} \qquad \beta_{ij} &= 4y_{ij} - 2y_i - 2y_j \end{aligned}\right\}$$

Substituting we find that the response for the ternary mixture is

$$y_{ijk} = (4/9)(y_{ij}+y_{ik}+y_{jk}) - (1/9)(y_i+y_j+y_k)$$

and the difference between the observed mean response of the ternary mixture and that predicted is given by

$$d_{ijk} = \hat{y}_{ijk} - (4/9)(\hat{y}_{ij}+\hat{y}_{ik}+\hat{y}_{jk}) + (1/9)(\hat{y}_i+\hat{y}_j+\hat{y}_k) \qquad (11)$$

where the carets denote mean values.

Suppose r_1 observations are made with the pure components, r_2 with the binary mixtures and r_3 with the ternary mixture. If the observations are independent with equal variance σ^2, then the variance of d_{ijk} is given by

$$\mathrm{Var}(d_{ijk}) = \sigma^2\left(\frac{1}{r_3} + \frac{16}{27}\,\frac{1}{r_2} + \frac{1}{27}\,\frac{1}{r_1}\right) \qquad (12)$$

The total number of observations $r_3+3r_2+3r_1$ is fixed and the minimum may be found using Lagrangian multipliers

$$L = \sigma^2\left(\frac{1}{r_3} + \frac{16}{27}\,\frac{1}{r_2} + \frac{1}{27}\,\frac{1}{r_1}\right) + \lambda\,(r_3+3r_2+3r_1)$$

Differentiating this respectively with respect to r_1, r_2 and r_3 and setting the derivatives equal to zero it is found that the minimum occurs when the observations are in the proportion

$r_1 : r_2 : r_3 = 1 : 4 : 9$.

For example, if we have 10-15 observations then to test a difference d_{ijk} we should take $r_1 = 1$, $r_2 = 2$, $r_3 = 4$ or 5. However, if one wished to place confidence limits on β_i, the coefficients of the polynomial associated with the pure components, one would have to use as a measure of the standard deviation a figure obtained from a consideration of the observations made at the binary, ternary etc., points as the contribution made by the pure components would be zero for a single observation. This is an unattractive arrangement and it is suggested that at least two observations are made at each point and the ratio 1 : 4 : 9 implemented as a secondary consideration.

The essential difference between this allocation of observations and that of the previous section is that in this case the observations are being allocated for the purpose of testing a goodness of fit while in the previous section they are allocated for the purpose of estimating the response curve after the type of polynomial had been chosen. A polynomial having the same number of coefficients as points in the lattice will fit the observed mean value at each lattice point exactly and no measure of lack of fit is obtained.

Testing goodness of fit.

To test the adequacy of the model we need to compare the calculated response with an observed response somewhere other than at one of the lattice points for the reasons given at the end of the last paragraph. For this reason it is necessary to introduce into the experiment observations taken in regions of particular interest, the previously referred to "check points". We have seen that in the case of a quadratic polynomial the centre

point may be used as a check point and introduced into the calculation to form the special cubic if the quadratic is found to be inadequate. For other models the experimenter has a free choice as to the position of his check points.

In order to judge the suitability of a model the student's t-value for the difference between the observed and calculated values needs to be determined.

For the quadratic model using the centre point of the simplex as a check point the test will be

$$t = d_{ijk} \Big/ \mathrm{Var}(d_{ijk})$$

where the numerator and denominator are given above. If the number of observations is the same at each lattice point this reduces to

$$t = d_{ijk} \cdot \frac{27}{44} \frac{\sqrt{r}}{\sigma}$$

where r is the number of replications at each lattice point.

When using points other than the centre of the simplex as check points the following formulae are used to compute the variance of the predicted values. The equation for the variance of the cubic has been derived earlier, but it is quoted again for completeness.

Quadratic Model.

$$\mathrm{Var}(\tilde{y}) = \sigma^2 \left(\sum_{1 \leq i \leq q} a_i^2 / r_i + \sum_{1 \leq i < j \leq q} a_{ij}^2 / r_{ij} \right)$$

where

$$a_i = x_i(2x_i - 1) \text{ and } a_{ij} = 4x_i x_j$$

Special Cubic Model.

$$\mathrm{Var}(\tilde{y}) = \sigma^2\left(\sum_{1\leq i\leq q} b_i^2/r_i + \sum_{1\leq i<j\leq q} b_{ij}^2/r_{ij} + \sum_{1\leq i<j<k\leq q} b_{ijk}^2/r_{ijk} \right)$$

where

$$b_i = (x_i/2)\left(6x_i^2 - 2x_i + 1 - 3\sum_{j=1}^{q} x_j^2\right)$$

$$b_{ij} = 4x_i x_j(3x_j - 2)$$

$$b_{ijk} = 27x_i x_j x_k$$

Cubic Model.

$$\mathrm{Var}(\tilde{y}) = \sigma^2\left(\sum_{1\leq i\leq q} C_i^2/r_i + \sum_{1\leq i<j\leq q} C_{iij}^2/r_{iij} + \sum_{1\leq i<j\leq q} C_{iij}^2/r_{ijj} + \sum_{1\leq i<j<k\leq q} C_{ijk}^2/r_{ijk} \right)$$

where

$$C_i = (x_i/2)(3x_i - 1)(3x_i - 2)$$

$$C_{iij} = (9/2)x_i x_j(3x_i - 1)$$

$$C_{ijj} = (9/2)x_i x_j(3x_j - 1$$

$$C_{ijk} = 27x_i x_j x_k$$

Notice that if the same number of observations are taken at each lattice point all of the above equations may be more compactly written

$$\mathrm{Var}(\tilde{y}) = \sigma^2\, z/r$$

where r is the number of observations at each point and z is the sum of the squares of the coefficients, a function which depends only on the composition of the mixture.

Lambrakis (1968a) has shown that a general expression for the coefficients, z, of the variance equations is

$$a_{Sr} = \prod_{j=1}^{k} \{(mx_j)!/r_j!(mx_j-r_j)!\}$$

where Sr is any subset of the r elements of $(1,2,\ldots,q)$, r_j is the number of occurrences of a letter or number in the subscript and m is the degree of polynomial to be fitted.

Example.

For the cubic model, $(q,3)$, $m = 3$

a) $C_1'(\equiv C_{111})$, hence $r_1 = 3$

$$C_{111} = \prod_{j=1}^{1} \{(3x_1)!/3!(3x_1-3)!\} = (x_1/2)(3x_1-1)(3x_1-2)$$

B) C_{112}, hence $r_1 = 2$ and $r_2 = 1$

$$C_{112} = \prod_{j=1}^{2} \{(3x_j)!/r_j!(3x_j-r_j)!\} = (9/2)x_1x_2(3x_1-1)$$

c) C_{123}, hence $r_1 = r_2 = r_3 = 1$

$$C_{123} = \prod_{j=1}^{3} \{(3x_j)!/1!(3x_j-1)!\} = 27x_1x_2x_3$$

Compare these with the expression on page 18

As the value of z depends only on composition its value can be represented by contour lines on a simplex. Gorman and Hinman have published four simplexes from the quadratic to the quartic model which show these contours. But the dimensions are so small (2" × 2") that they can be used only for the coarsest measurements. To remedy this a table of z values has been calculated at intervals of 1.0(0.02)0.34 and is given in Appendix 4 The function is symmetrical within the area bounded by the lines joined the centre of one side, the centroid and an apex. Hence the above intervals cover all cases.

Synergism and Antagonism.

The polynomials discussed can conveniently be re-

garded as being composed of two major parts; a linear part and a non-linear part. The first summation represents a linear combination of the mixture variables such as would occur if the density of a mixture of three liquids which underwent no volume change on mixing were measured. The second summation is the non-linear part and represents deviations from linearity which may be either positive or negative. If the deviation is negative the effect is said to be antagonistic; if positive the effect is said to be synergistic a term introduced in this context by Macht (1929). Sometimes agonistic is used as synonymous with synergistic. If the terms are antagonistic they will deflate the response below what one would expect from a linear model and vice-versa in the case of synergistic terms.

The non-linear part is sometimes sub-divided to give binary synergistic (antagonistic) terms, ternary synergistic (antagonistic) terms etc.,

There has been some attempt to identify these terms with Yates's idea of interaction effects (Quenouille 1959). But Scheffe (1961) strongly rejects this and studiously avoids the term "interaction". He argues that for, say, a three component mixture the experimental points needed to define a two-factor interaction cannot be chosen without varying the third factor in such a way as to produce somehow an "interaction" of all three factors. The usual notions seem not to be applicable in this case.

A knowledge of the synergism and antagonism can obviously be useful in giving some indication of the underlying mechanism and it has been suggested that it might help in the determination of Refutas blending numbers. These are numbers by which the viscosity of a mixture of oils can be precalculated, as the viscosity of a mixture of oils is not a linear function of the

viscosity of the individual components.

In this section it seems appropriate to mention a piece of jargon sometimes encountered in this context; i.e. isobols. These are lines representing the combination of materials just necessary to cause a standard effect (Loewe,1928).

Response Surfaces.

Because of the restriction that the proportions of a mixture must sum to unity a three component mixture may be represented by a point on triangular graph paper, while a four component mixture may be represented by a point in a regular tetrahedron. Hence a response surface may be plotted.

Example.

The following example will illustrate the application of the previous theory to a problem in petrol blending which is a modification of the one in Gorman and Hinman (1962) concerning the octane rating of petrol blends.

In practice the octane rating of different hydrocarbons show marked deviations from linearity when blended together. Paraffins with paraffins and olefins with olefins are sensibly linear while blends of paraffins with olefins or olefins with aromatics are not. With pure hydrocarbons these deviations may be quite large. But with commerical petrols, which contain a large number of hydrocarbons, deviations may be quite small. The reason for these deviations are probably associated with the different oxidation mechanisms of the different hydrocarbons.

The data in Table 2 represents octane ratings of blends of 3 different grades of petrol. It was decided to fit a quadratic model to the data or, if that was found to be not adequate, to fit a special cubic. The centroid corresponding to

the mixture $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$ was to be used as the check point for testing the quadratic. Hence, the observations were taken in the ratio given. A further observations was also made at the point (0.295, 0.405, 0.300).

Table 2 : Octane Ratings of Blends of Petrol.

Blend	Volume Fraction X_1	X_2	X_3	Observed Octane No.	Mean	
1	1	0	0	100.9	100.90	y_1
2	0	1	0	85.4	85.4	y_2
3	0	0	1	85.5	85.4	y_3
4	$\frac{1}{2}$	$\frac{1}{2}$	0	88.8 89.3	89.05	y_{12}
5	$\frac{1}{2}$	0	$\frac{1}{2}$	99.3 90.7	95.5	y_{13}
6	0	$\frac{1}{2}$	$\frac{1}{2}$	85.5 85.4	85.45	y_{23}
7	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	88.3 88.5 88.9 88.5	88.66	y_{123}
8	0.295	0.405	0.300	87.0 88.0	87.5	

$S^2 = 0.295$

7d. of f.

Calculation of Coefficients of the Quadratic Model.

$\beta_1 = y_1 = 100.9$

$\beta_2 = y_2 = 85.4$

$\beta_3 = y_3 = 85.5$

$\beta_{12} = 4y_{12}-2y_1-2y_2 = 4(89.05)-2(100.90)-2(85.40) = -16.4$

$\beta_{13} = 4y_{13}-2y_1-2y_3 = 4(90.05)-2(100.90)-2(85.50) = -10.8$

$\beta_{23} = 4y_{23}-2y_2-2y_3 = 4(85.45)-2(85.40)-2(85.50) = 0.0$

The quadratic model then is

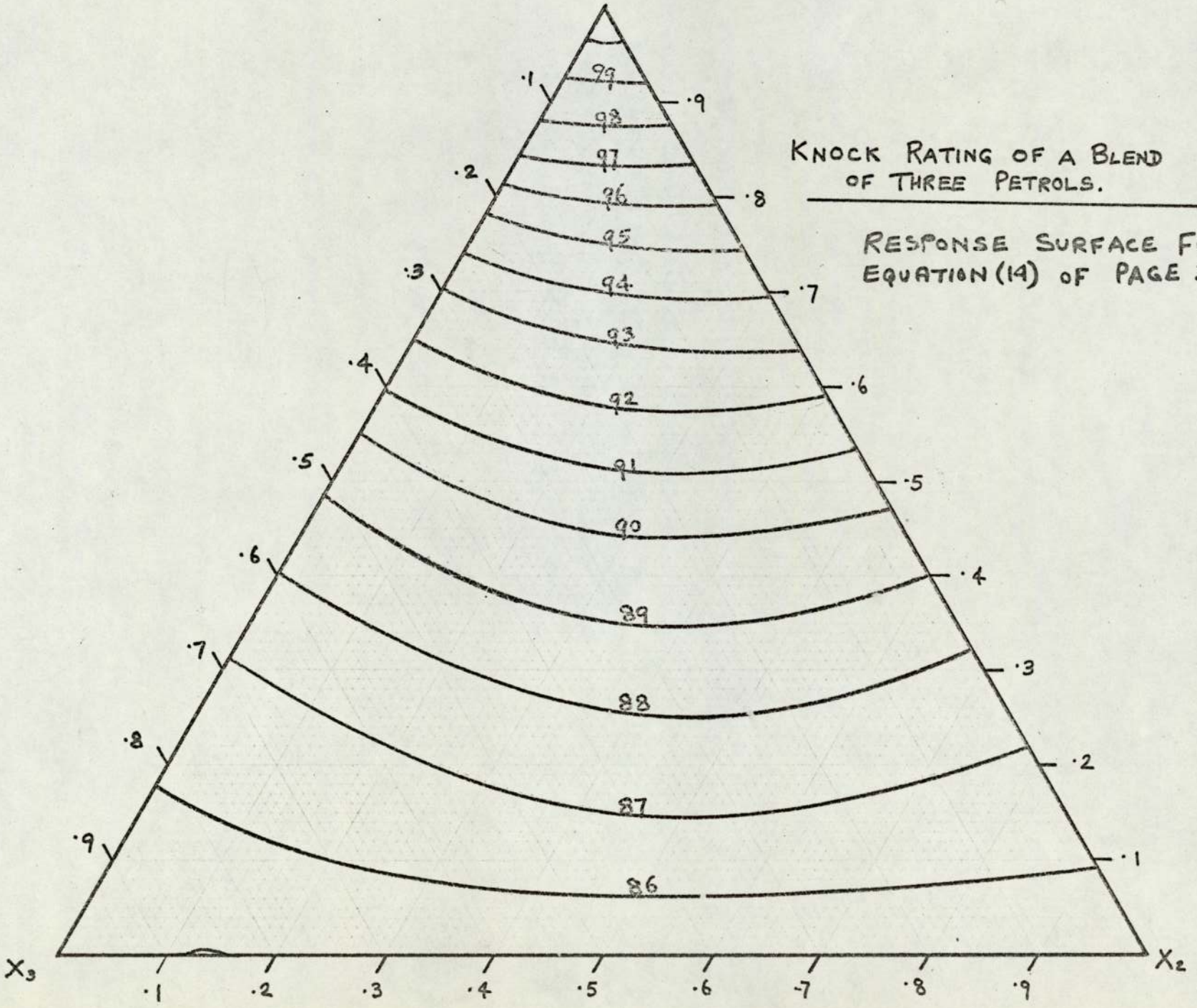
$$y = 100.9x_1+85.4x_2+85.5x_3 - 16.4x_1x_2 - 10.8x_1x_3 \qquad (13)$$

Using this equation the response at the centre point where $x_1 = x_2 = x_3 = \frac{1}{3}$ may be calculated

$$y_{123} = 100.9(\tfrac{1}{3}) + 85.4(\tfrac{1}{3}) + 85.5(\tfrac{1}{3}) - 16.4(\tfrac{1}{9}) - 10.8(\tfrac{1}{9})$$

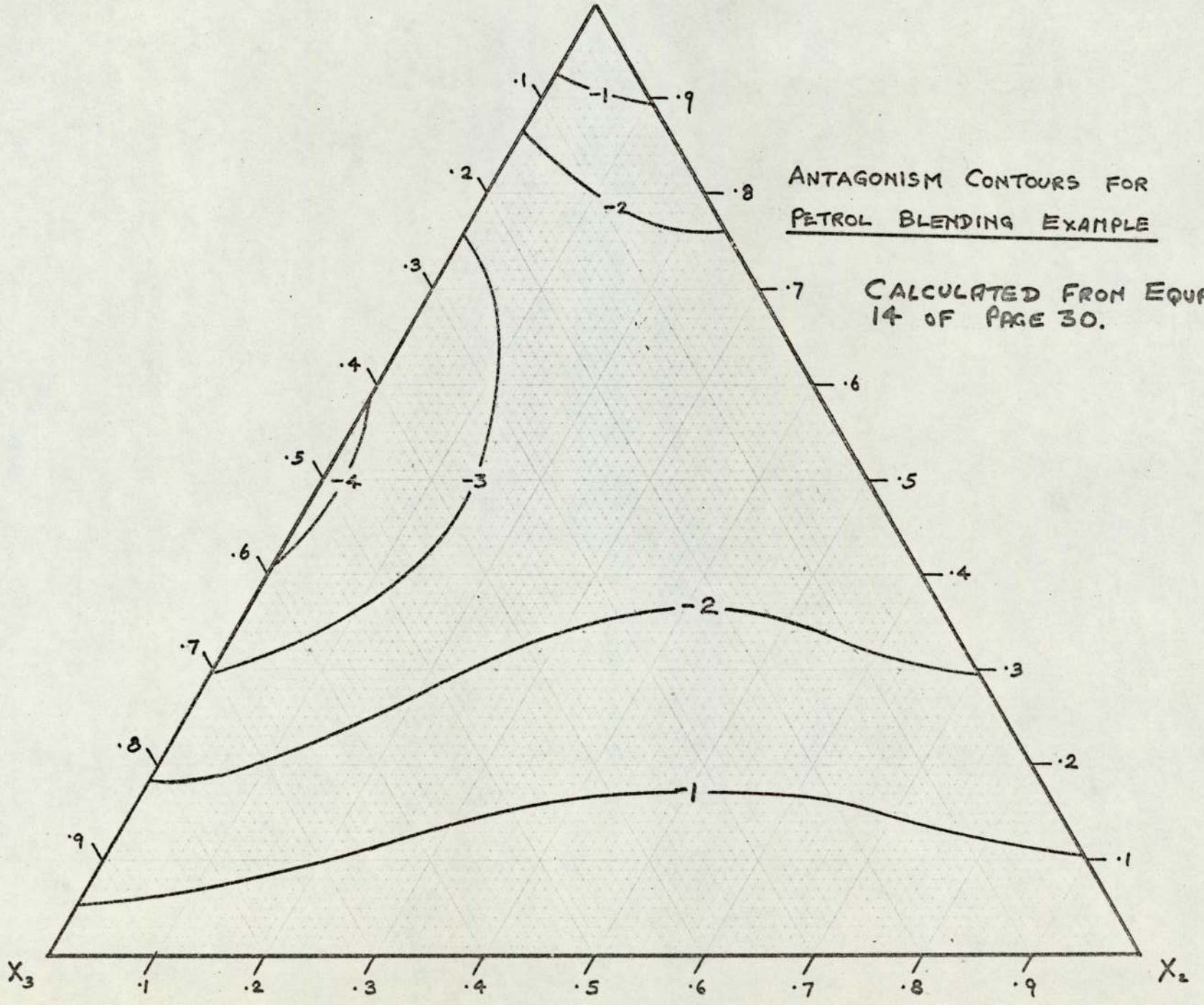
$$= 87.58$$

By (11), $d_{123} = 88.6 - 87.58 = 1.08$ and by (12)



KNOCK RATING OF A BLEND OF THREE PETROLS.

RESPONSE SURFACE FROM EQUATION (14) OF PAGE 30.


ANTAGONISM CONTOURS FOR
PETROL BLENDING EXAMPLE
CALCULATED FROM EQUATION
14 OF PAGE 30.
-1
-2
-3
-4
-2
-1
X3
X2
·1
·2
·3
·4
·5
·6
·7
·8
·9

$$\text{Var}(d_{123}) = 0.295\left(\tfrac{1}{3} + \tfrac{16}{27}\cdot\tfrac{1}{2} + \tfrac{1}{27}\right) = 0.197$$

and $t = 1.08/\sqrt{0.197} = 2.53$.

The value of t is significant at the 5% level with 7 degrees of freedom ($t_{0.025,7} = 2.36$). Further a difference of 1.08 is more than would be acceptable in practice. Therefore, a more complex model is necessary and we augment (13) by the term $\beta_{123}x_1x_2x_3$, where β_{123} for the special cubic is given by (9)

$$\beta_{123} = 27(88.6) - 12(89.05+85.5+85.45) + 3(100.9+85.4+85.5) = 29.22$$

and the model now is

$$y = 100.9x_1+85.4x_2+85.5x_3-16.4x_1x_2-10.8x_1x_3+29.2x_1x_2x_3 \qquad (14)$$

We use the checkpoint, blend 8, to test this new model. The observed value at the point (0.295,0.405,0.3) is 87.5 and the value calculated from the special cubic model above is 87.3, a difference 0.2 and we now have to calculate the standard error in order to carry out a t-test on this difference. We can calculate the b coefficients using the x_1,x_2,x_3 coordinates of blend 8

b_1	= -0.0134	b_1^2	= 0.000180
b_2	= 0.0306	b_2^2	= 0.000935
b_3	= 0.0125	b_3^2	= 0.000156
b_{12}	= 0.0478	b_{12}^2	= 0.002284
b_{13}	= -0.0761	b_{13}^2	= 0.005793
b_{23}	= 0.0559	b_{23}^2	= 0.003124
b_{123}	= 0.9677	b_{123}^2	= 0.936443

E.g. $b_1 = (0.295/2)(6\times0.295^2-2\times0.295+1 - 3(0.295^2+0.405^2+0.3^2))$

$= -0.0134$

We can now find the variance of the predicted value $\text{Var}(\tilde{y})$

$$\mathrm{Var}(\tilde{y}) = 0.295\left(\frac{0.000180}{1} + \frac{0.000935}{1} + \frac{0.000156}{1}\right.$$

$$+ \frac{0.002284}{2} + \frac{0.005793}{2} + \frac{0.003124}{2}$$

$$\left. + \frac{0.936443}{4}\right) = 0.071$$

The variance of the difference between the observed and predicted values is the sum of the variances of the observed value (0.197) and the predicted value (0.071). Hence, the t-test is

$$t = \frac{87.50 - 87.29}{\sqrt{(0.197+0.071)}} = 0.512$$

Entering the t-table with 7 degrees of freedom we find that the result is not significant at the 50% level ($t_{0.25,7} = 0.711$) and we conclude that the model is adequate.

If several points in composition are to be used to check the lack of fit a t value is calculated for each. However, these t values are referred to the $\alpha/(2k)$ point instead of the usual $\alpha/2$ point of the t distribution where k is the number of check points in composition. This will be amplified in the next section. Meanwhile, we will notice that this usually leads to untabulated figures for the deviate and interpolation is usually necessary using, for instance Federighi's (1959) or Fisher's (1925) detailed t-tables. From Fisher's tables a useful graph of t against percentage point may be plotted for various degrees of freedom. Alternatively, Dunn (1959) gives an abbreviated table of the $1-0.05/2k$ point of the Students t-distribution from which intermediate value may be obtained by interpolating with the reciprocal of the degrees of freedom.

Precision of Estimates.

Dunn (1959,1968) has considered the problem of finding confidence intervals for the means of dependent normally distributed

variables when nothing is known about the correlation which might exist amongst them. There are occasions when one does not wish to make the laborious calculations to obtain the confidence band for a regression curve but rather to look for separate confidence intervals for each mean with the assurance that, with high probability, each interval of the set contains its mean.

In the case of mixtures, the several different observations taken over the simplex are made on mixtures of the same materials (some in zero proportion) and the data may be correlated and the degree of correlation is unknown. In this case Dunn has shown that, if only an estimate of the variances is known the confidence intervals with confidence level $1-\alpha$ of the k sample means are

$$\bar{y}_i \pm C_\alpha S/\sqrt{n} \qquad (1 \leq i \leq k)$$

where $\bar{y}_i$ is the mean of the i th. set of observations and C_α is the $1 - (\alpha/2k)$ point of the Students t-distribution with $k(n-1)$ degrees of freedom and S is the overall variance of k observations replicated n times.

Now, when testing lack of fit using k check points one is, in effect, carrying out a joint test at k points in compositions - in essence two linear combinations of all the observations. This is the basis last paragraph of the previous section.

There is, however, another manipulation, which will be illustrated, in which we find the confidence limits at <u>any</u> point on the simplex and for this we use the t-value referred to $\alpha/(2k)$ point where k is the number of regression coefficients in the equation. The argument here is that we are concerned with simultaneously all confidence intervals which might be constructed for all compositions, each of which however concerns

linear combinations of the k regression coefficients.

Let us assume for the sake of the exercise that we have derived (13) on the basis of 3 observations at each lattice point, with no check points and that the variance is 0.36. The number of degrees of freedom is $(3 \times 7)-7 = 14$.

Suppose we now wish to calculate 95% confidence limits at the point $x_1 = 0.56$, $x_2 = 0.34$, $x_3 = 0.10$ at which the octane rating is 90.9. For 95% confidence limits we need to find δ such that

$$Pr(y - \delta < y < y + \delta) = 0.95$$

where

$$\delta = t_{\alpha/2k,f}\sqrt{Var(y)}.$$

t is taken at the $\alpha/2k$ level because of the k parameters in the regression equation.

We have then

α	=	0.05
k	=	7, no. of constants in model
f	=	14 degrees of freedom
$\alpha/2k$	=	0.00357
t	=	3.11 at 0.00357 level for 14 d. of f.
z	=	0.5757 at $x_1=0.56$, $x_2= 0.34$, $x_3=0.10$ from tables in Appendix 4
r	=	3 observations at each lattice point
S^2	=	0.36

Hence, $\sqrt{Var(y)} = \sqrt{\sigma^2 z/r} = \sqrt{0.36 \times 0.5757/2} = 0.32$

and $\delta = 3.11 \times 0.32 = 0.99$

Thus the octane rating lies between 90.9 ± 0.99 i.e. 89.9 to 91.9.

Simplex Centroid Designs.

A (q,m) simplex lattice gives an equally spaced distribution of points over the simplex and has just enough points to enable a polynomial of degree m to be uniquely fitted.

An objection to the simplex lattice design is the following: suppose we wish to predict the response of a 4 component mixture. We would use a (4,m) simplex lattice. If m = 2 we are using observations on pure components only; if m = 3 we are using observations on pure components and binary and ternary mixtures. Only when m = 4 do we have any observations of the effect of quaternary mixtures which is the real object of the experiment. Further for m > 2 the simplex lattice contains components in unequal proportions (see Fig.2).

The simplex centroid designs differ in that they contain observations on mixtures consisting of every subset of the composition in equal proportions. This corresponds to the points $(x_1, x_2, \ldots, x_q)$ of the simplex obtained by taking the q permutations of $(1,0,\ldots,0)$, the $\binom{q}{2}$ permutations of $(\frac{1}{2},\frac{1}{2},0,\ldots,0)$, the $\binom{q}{3}$ permutations of $(\frac{1}{3},\frac{1}{3},\frac{1}{3},0,\ldots,0)$ etc. and the point $(1/q, 1/q, 1/q \ldots, 1/q)$.

Simplex centroid designs are also computationally simpler than the lattice designs and, as will be shown, can be developed to a more sophisticated level. One has already been met, the special cubic modal (Fig. 2).

Calculation of Coefficients of the Simplex Centroid Models.

A polynomial which has as many coefficients as there are points in the centroid design is

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j + \sum_{1\leq i<j<k\leq q} \beta_{ijk} x_i x_j x_k + \cdots$$

$$+ \beta_{12\ldots q}\, x_1 x_2 \cdots x_q \qquad (15)$$

Let the response of the pure component be y_i. Putting $x_i = 1$ (and hence $x_j = x_k = \ldots = x_q = 0$) and inserting this value we get

$$\beta_i = y_i$$

Let the response of the binary mixture b y_{ij}. Putting $x_i = x_j = \frac{1}{2}$ (and hence $x_k = \ldots = x_q = 0$) and inserting this value we get

$$y_{ij} = \tfrac{1}{2}\beta_i + \tfrac{1}{2}\beta_j + \tfrac{1}{4}\beta_{ij}$$

which gives

$$\beta_{ij} = 2(2y_{ij} - 1(y_i + y_j))$$

Let the response of the ternary mixture b y_{ijk}. Putting $x_i = x_j = x_k = \frac{1}{3}$ (rest = 0) and inserting this value we get

$$\beta_{ijk} = 3(3^2 y_{ijk} - 2^2(y_{ij}+y_{ik}+y_{jk}) + 1^2(y_i+y_j+y_k)$$

These equations seem to indicate a pattern and Scheffe has shown that the general formula is given by

$$\beta_{S_r} = r(r^{r-1}\,\phi_r(S_r) - (r-1)^{r-1}\phi_{r-1}(S_r) + (r-2)^{r-1}\phi_{r-2}(S_r) +$$

$$\ldots\ldots + (-1)^{r-1}\,1^{r-1}\,\phi_1(S_r))$$

where S_r is any subset of r elements of $(1,2,\ldots,q)$ and $\phi_t(S_r)$ is the sum of all the responses of the t-nary mixtures with equal proportions from the r components in S_r.

Example.

We will calculate the next β coefficients, β_{ijkm}.

$$r = 4$$

$$\beta_{S_4} = 4(4^3\phi_4(S_4)-3^3\phi_3(S_4)+2^3\phi_2(S_4)-1^3\phi_1(S_4)) \qquad (16)$$

Now $S_4 = (ijkm)$

and $\phi_4(ijkm) = y_{ijkm}$

$$\phi_3(ijkm) = y_{ijk}+y_{ijm}+y_{ikm}+y_{jkm}$$

$$\phi_2(ijkm) = y_i+y_j+y_k+y_m$$

Hence

$$\beta_{ijkm} = 4(4^3y_{ijkm}-3^3(y_{ijk}+y_{ijm}+y_{ikm}+y_{ikm}) +2^3(y_{ij}+y_{ij}+y_{im}+y_{jk}+y_{jm}+y_{km}) -1^3(y_i+y_j+y_k+y_m))$$

It can be appreciated that the coefficients are much more easily derived than in the case of the simplex lattice designs.

In the case of what was previously called the special cubic model the distribution of points over the simplex is identical with that of the simplest centroid design i.e. $(1,0,0)$, $(0,1,0)$, $(0,0,1)$, $(\frac{1}{2},\frac{1}{2},0)$,$(\frac{1}{2},0,\frac{1}{2})$, $(0,\frac{1}{2},\frac{1}{2})$, $(\frac{1}{3},\frac{1}{3},\frac{1}{3})$, a comparison of the above shows that they are identical.

Variance of Predicted Responses.

To illustrate the method by which the variance of the simplex centroid designs are obtained we will calculate that for $m = 3$

$$y = \beta_1x_1 + \beta_2x_2 + \beta_3x_3 + \beta_{12}x_1x_2 + \beta_{13}x_1x_3 + \beta_{23}x_2x_3 + \beta_{123}x_1x_2x_3$$

If in this formula we replace the β's by their estimates evaluated in terms of the responses we may collect coefficients of like

terms to obtain the predicted response $\tilde{y}$, where

$$\tilde{y} = \sum_{1\leqslant i\leqslant 3} a_i \hat{y}_i + \sum_{1\leqslant i<j\leqslant 3} a_{ij}\hat{y}_{ij} + a_{123}\,\hat{y}_{123}$$

whence the variance is given by

$$\mathrm{Var}(\tilde{y}) = \sigma^2\left(\sum_{1\leqslant i\leqslant 3} a_i^2/r_i + \sum_{1\leqslant i<j\leqslant 3} a_{ij}^2/r_{ij} + a_{123}^2/r_{123} \right)$$

where r is the number of observations at each point.

Let a_1 be the coefficient associated with y_1. Then if one carries out the process outlined above one finds that

$$a_1 = x_1 + 2\sum_{j=2}^{3} x_1x_j + 3\sum_{2\leqslant j<k\leqslant 3} x_1x_jx_k$$

$$= x_1\left(1 - 2\sum_{j=2}^{3} x_j + 3\sum_{2\leqslant j<k\leqslant 3} x_jx_k\right)$$

The last term is merely $3x_1x_2x_3$, but it is written in this form to help in the simplification. In order to simplify we have to utilise the fact that

$$\sum_{2\leqslant j<k\leqslant q} x_jx_k = \tfrac{1}{2}\left\{ \left(\sum_{j=2}^{q} x_j^2 \right) - \sum_{j=2}^{q} x_j^2 \right\}$$

This can be shown by the following expansion

$$\begin{aligned}(x_2 + x_3 + \ldots + x_q)^2 = {} & (x_2^2 + x_3^2 + \ldots + x_q^2) \\ & + 2x_2(x_3 + \ldots + x_q) \\ & + 2x_3(x_4 + \ldots + x_q) \\ & + \ldots\ldots\ldots \\ & + 2x_{q-1}x_q\end{aligned}$$

$$\text{i.e.} \quad \left(\sum_{j=2}^{q} x_j\right)^2 = \sum_{j=2}^{q} x_j^2 + 2 \sum_{2 \leq j < k \leq q} x_j x_k$$

whence the result follows.

$$\text{Using} \quad \sum_{j=2}^{q} x_j = 1 - x_1, \text{ and setting } q = 3$$

simplifies to

$$a_1 = (x_1/2)\left(3x_1^2 - 2x_1 + 1 - 3\sum_{j=2}^{3} x_j^2\right)$$

Similarly

$$a_{12} = 4x_1x_2 - 12x_1x_2x_3 = 4x_1x_2(3x_1 + 3x_2 - 2)$$

and

$$a_{123} = 27x_1x_2x_3$$

Process Variables.

Suppose one wishes to examine the mixture variables at different levels of other factors which we will call process variables. E.g. the effect of mixtures of feeding stuffs on two different breeds of cow when given for two different periods of time or the road octane number of a blend of petrols when the make and speed of the car are varied. If there were no mixture variables the mathematical model corresponding to three factors A,B and C varied at two levels can be expressed as

$$y_{ijk} = I + A_i + B_j + C_k + A_iB_j + A_iC_k + B_jC_k + A_iB_jC_k + e_{ijk}$$

where y_{ijk} = the response with A,B and C at the ith,jth and kth level ($i = \pm 1, j = \pm 1, k = \pm 1$)

I = the true mean of all the trials

A_i = the true mean of all the trials in which A is at its ith level

B_j and C_k are similarly defined

A_iB_j = the interaction of A and B with A at the ith level and B at the jth level

A_iC_k, B_jC_k and $A_iB_jC_k$ are similarly defined.

To include process variables in the centroid design we carry out a complete factorial experiment at each lattice point and determine the responses in the form given on the next page. These variable responses involving A, B, C etc. are used to calculate variable coefficients involving A,B,C etc. by using the formulae derived from (16) and we get a regression equation which can be solved for all values of the mixture variables at the two levels of the process variables.

TABLE 3 : Responses of 3 mixture variables and 2 process variables.

$x_1 x_2 x_3$	A	B	Response	Response function
	−	−	100	
100	+	−	108	$y_1 = 150.25+7.25A+46.25B-3.25AB$
	−	+	186	
	+	+	207	
	−	−	84	
010	+	−	194	$y_2 = 160.25+19.25A+21.25B-35.75AB$
	−	+	198	
	+	+	165	
	−	−	85	
001	+	−	98	$y_3 = 117.25+10.25A+25.75B+3.75AB$
	−	+	129	
	+	+	157	
	−	−	89	
$\frac{1}{2}\frac{1}{2}0$	+	−	202	$y_{12} = 139.75+42.75A-5.75B-13.75AB$
	−	+	105	
	+	+	163	
	−	−	90	
$\frac{1}{2}0\frac{1}{2}$	+	−	120	$y_{13} = 132.5+122.5A+27.5B+7.5AB$
	−	+	130	
	+	+	190	
	−	−	85	
$0\frac{1}{2}\frac{1}{2}$	+	−	140	$y_{23} = 170.0+27.5A+57.5B$
	−	+	200	
	+	+	255	
	−	−	88	
$\frac{1}{3}\frac{1}{3}\frac{1}{3}$	+	−	144	$y_{123} = 204.75+12.75A+88.75B-15.25AB$
	−	+	296	
	+	+	291	

Example.

It is desired to investigate the effect of a mixture of 3 components at two levels of the factors A and B. The observed responses are shown in Table 3.

The β coefficients corresponding to the responses of the pure components are

$$\beta_1 = 150.25 + 7.25A + 46.25B - 3.25AB$$

$$\beta_2 = 160.25 + 19.25A + 21.25B - 35.75AB$$

$$\beta_3 = 117.25 + 10.25A + 25.75B + 3.75AB$$

The variable coefficients for the binary and ternary responses are obtained by inserting the values of the variable responses.

E.G. $\beta_{12} = 4y_{12} - 2y_1 - 2y_2$

$$= 4(139.75+42.75A-5.75B-13.75AB)$$
$$-2(150.25+7.25A+46.25B+3.25AB)$$
$$-2(160.25+19.25A+21.25B-35.75AB)$$

$$= -62.0+118.0A-150.0B + 10.0AB$$

Similarly

$$\beta_{13} = -5.0 + 55.0A - 34.0B + 16.0AB$$

$$\beta_{23} = 125.0 + 51.0A + 136.0B + 64.0AB$$

$$\beta_{123} = 1504.5-858.5A+1725.0B-442.5AB$$

These seven variable coefficients are inserted in (15) with $q = 3$ to give an equation for the responses involving not only the mixture variables but also the process variables A and B which can take values -1 to +1. An alternative method of calculating these coefficients is given in Appendix 3.

$$
\begin{aligned}
y = {} & (150.25 + 7.25A + 46.25B - 3.25AB)x_1 \\
& + (160.25 + 19.25A + 21.25B - 35.75AB)x_2 \\
& + (117.25 + 10.25A + 25.75B + 3.75AB)x_3 \\
& + (-62.0 + 118.0A - 150.0B + 10.0AB)x_1x_2 \\
& + (-5.0 + 55.0A - 34.0B + 16.0AB)x_1x_3 \\
& + (125.0 + 51.0A + 136.0B + 64.0AB)x_2x_3 \\
& + (1504.5 - 858.5A + 172.0B - 442.5AB)x_1x_2x_3
\end{aligned}
$$

The response contours and the synergism and antagonism of the above equation for all combinations of levels of A and B are plotted in Figs. 5 and 6.

At the higher level of B the pattern and values of the responses do not change a great deal in passing from the higher level of A to the lower level. The maximum for A = +1, B = +1 is 287.5 and that for A = -1, B = +1 is 307.0. Similarly the synergism and antagonism patterns do not alter markedly. However, there is a considerable difference in the response surfaces for B = -1 whether compared with each other or with the A = +1 surfaces. Of the four surfaces that for A = +1 and B = -1 is the odd one out and so is its synergism/antagonism plot.

The following conclusions can be drawn from these surfaces

1) If the highest value is required this should be taken at $x_1 = 0.28$, $x_2 = 0.36$, $x_3 = 0.36$ at the lower level of A and the upper level of B to give a value of 307.0.
2) If the lowest value is required this should be taken at $x_3 = 1.0$, $x_1 = x_2 = 0.0$ at the lower level of A and of B to give a value of 84.0.
3) If a value of, say, 200.0 is required there is an infinite choice except at the lower level of A and B.

Fig 5. Response Surfaces with Process Variables.

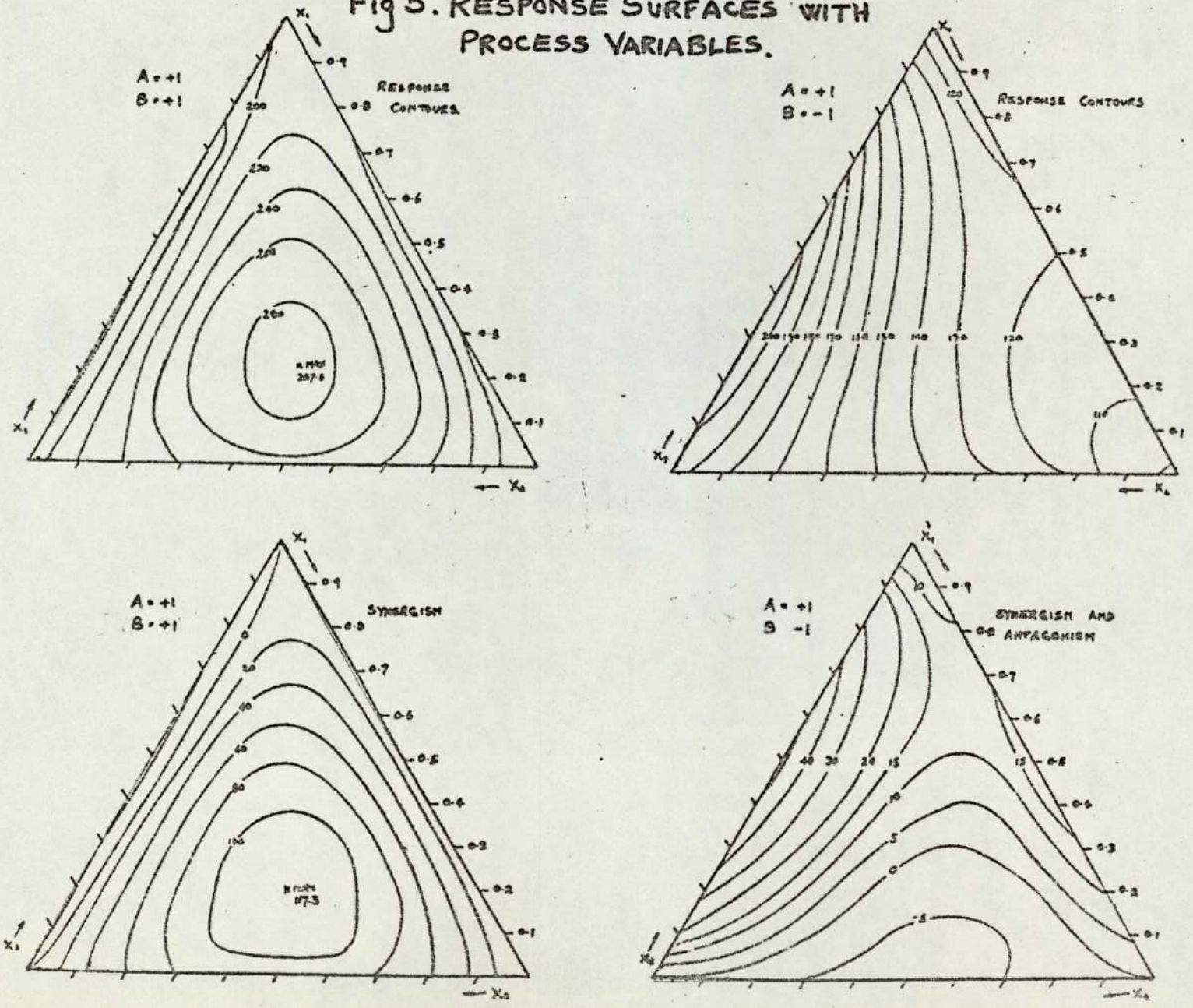
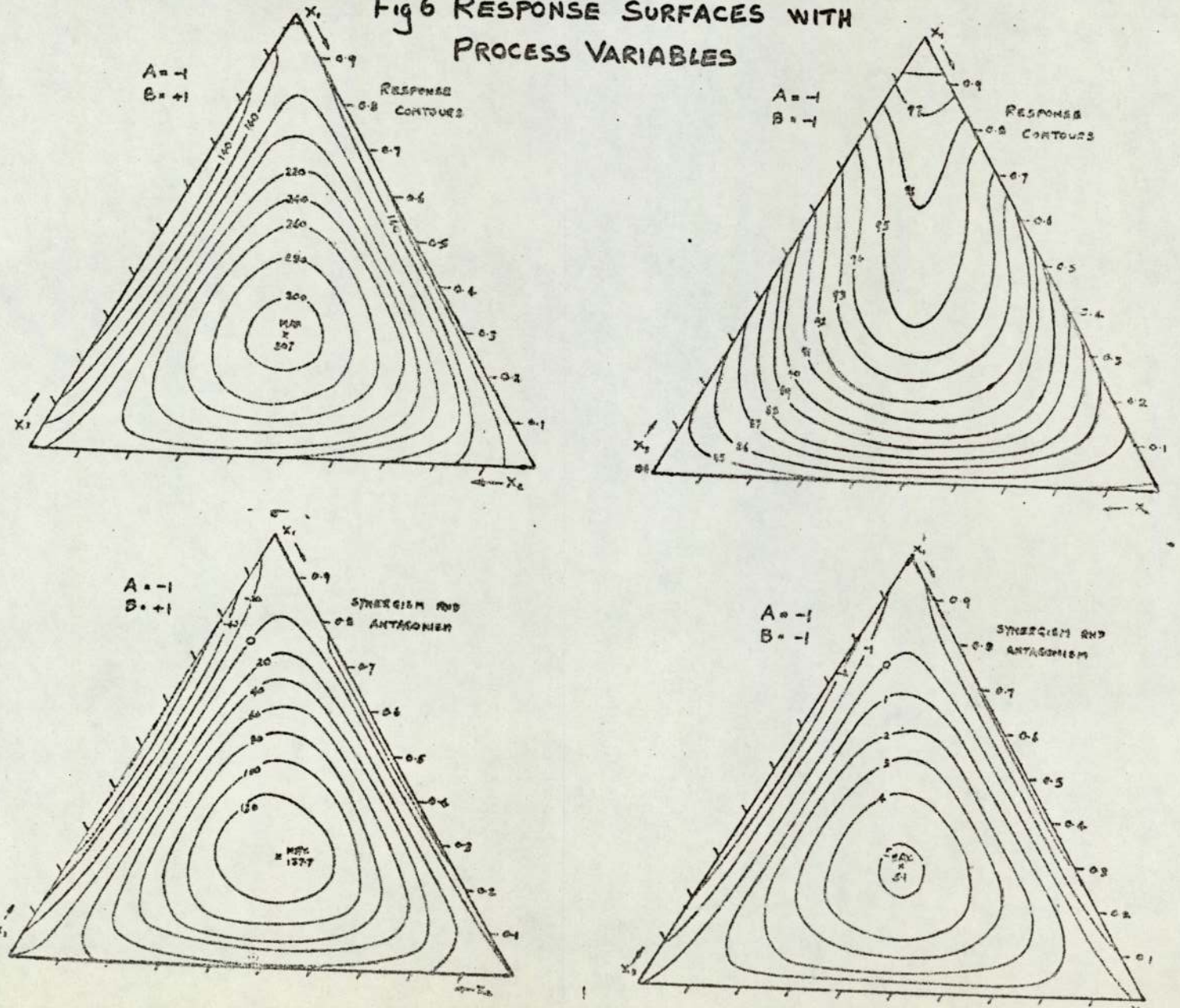


Fig 6 Response Surfaces with Process Variables



Variances and co-variances of Simplex Centroid Designs.

We shall assume that the various observations are normally distributed with a variance of σ^2. As $\beta_i = y_i$, var $\beta_i = \sigma^2$. Now, $\beta_{ij} = 4y_{ij} - 2y_i - 2y_j$; hence $\text{var}\beta_{ij} = 16\sigma^2 + 4\sigma^2 + 4\sigma^2 = 24\sigma^2$. Similarly $\text{var}\beta_{ijk} = 729\sigma^2 + 332\sigma^2 + 27\sigma^2 = 1188\sigma^2$ and so on.

All the expressions for the regression coefficients, apart from those for pure components, contain mixed terms; hence they are correlated and some of their co-variances and correlation coefficients will now be calculated. Taking, as an example, the terms for β_i and β_{ij} we have

$$\begin{aligned}
\text{covar}(\beta_i, \beta_{ij}) &= E(\beta_i \beta_{ij}) - E(\beta_i)\, E(\beta_j) \\
&= E(y_i[4y_{ij} - 2y_i - 2y_j]) \\
&\quad - E(y_i)\, E(4y_{ij} - 2y_i - 2y_j) \\
&= 4E(y_i, y_{ij}) - 4E(y_i)\, E(y_{ij}) \\
&\quad - 2\{E(y_i^2) - E(y_i)\, E(y_i)\} \\
&\quad - 2\{E(y_i, y_j) - E(y_i)(y_j)\} \\
&= 4\,\text{covar}(y_i, y_{ij}) - 2\text{var}(y_i) - 2\text{covar}(y_i, y_j)
\end{aligned}$$

As the observations are independent $\text{covar}(y_i, y_{ij}) = 0$ and $\text{covar}(y_i, y_j) = 0$. Hence, $\text{covar}(\beta_i, \beta_{ij}) = -2\text{var}(y_i) = -2\sigma^2$. The correlation coefficient is given by

$$\rho = \frac{\text{Covar}(\beta_i, \beta_j)}{\sqrt{\{\text{Var}(\beta_i)\ \text{Var}(\beta_j)\}}} = \frac{-2\sigma^2}{\sqrt{(\sigma^2\ 24\sigma^2)}} = -\frac{1}{\sqrt{6}}$$

Other variances and covariances can similarly be calculated. However, Scheffe has shown by using the theory of sets that the covariances of any pair of coefficients which we will call β and β' are given by

$$\text{covar}(\beta,\beta') = \sigma^2 \sum_{t=1}^{h} \binom{h}{t} r(-1)^{r-t}\, t^{r-1}\, r'(-1)^{r'-t}\, t^{r'-1} \qquad (17)$$

where r is the number of elements of the q mixture variables and h is the number of subscripts which the β's have in common. Where β and β' are identical $r = r' = h$ and (16) reduces to

$$\text{var}(\beta) = \sigma^2 r^2 \sum_{t=1}^{h} \binom{r}{t} t^{2r-2} = \sigma^2\, g(r) \qquad (18)$$

When there are n process variables present

$$\text{var}(\beta) = \sigma^2 r^2 \sum_{t=1}^{h} \binom{r}{t} t^{2r-2}/2^n = \sigma^2\, g(r)/n \qquad (19)$$

The manipulation of (17) and (18) will now be illustrated

Case 1

No elements in common i.e. β_i, β_j, β_k, etc.

$h = 0$ and $\text{covar}(\beta,\beta') = 0$.

Case 2

(a) One element in common i.e. β_i, β_{ij}, $r = 1$, $r' = 2$ and $h = 1$.

$$\text{covar}(\beta_i,\beta_{ij}) = \sigma^2 \sum_{t=1}^{1} \binom{1}{t} .1.(-1)^{1-t}\, t^{1-1}\, 2\, (-1)^{2-t}\, t^{2-1} = -2\sigma^2$$

which is identical with the result calculated earlier.

(b) One element in common i.e. β_{ij}, β_{ik}, $r = 2$, $r' = 2$, $h = 1$

$$\text{covar}(\beta_{ij},\beta_{ik}) = \sigma^2 \sum_{t=1}^{1} \binom{1}{t} .2.(-1)^{2-t}\, t^{1-1}\, 2\, (-1)^{2-t}\, t^{2-1} = 4\sigma^2$$

As an example of the manipulation of (18), take the case of a ternary mixture for which $r = 3$

$$\mathrm{Var}(\beta) = \sigma^2\, 9 \sum_{t=1}^{3} \binom{3}{t} t^{2\cdot 3-2}$$

$$= \sigma^2\, 9\left\{ \binom{3}{1} 1^4 + \binom{3}{2} 2^4 + \binom{3}{3} 3^4 \right\} = 1188\sigma^2$$

These equations enable covariance and correlation matrices to be easily drawn up.

Covariance matrix for β's up to β_{ijk}

	β_i	β_j	β_k	β_{ij}	β_{ik}	β_{jk}	β_{ijk}
β_i	1	0	0	-2	-2	0	3
β_j		1	0	-2	0	-2	3
β_k			1	0	-2	-2	3
β_{ij}				1	4	4	-36
β_{ik}					1	4	-36
β_{jk}						1	-36
β_{ijk}							1

All $\times \sigma^2$

Correlation Matrix for β's up to β_{ijk}

	β_i	β_j	β_k	β_{ij}	β_{ik}	β_{jk}	β_{ijk}
β_i	1	0	0	$-1/\sqrt{6}$	$-1/\sqrt{6}$	0	$1/6\sqrt{33}$
β_j		1	0	$-1/\sqrt{6}$	0	$-1/\sqrt{6}$	$1/6\sqrt{33}$
β_k			1	0	$-1/\sqrt{6}$	$-1/\sqrt{6}$	$1/6\sqrt{33}$
β_{ij}				1	1/6	1/6	-4/297
β_{ik}					1	1/6	-4/297
β_{jk}						1	-4/297
β_{ijk}							1

The value of g(r) given by (17) increases rapidly with r, being in excess of 10^7 for r = 5. (The actual value is 19,662,000). However the standard deviation obtained from (18) and (19) has to be multiplied by the variables $x_i, x_i x_j, x_i x_j x_k$ etc. and these have maximum

values when the variables within a set have equal values, i.e. for single components, 1; for binary mixtures, $\frac{1}{2}$; etc. and in general the maximum value over the simplex is r^{-r}. Hence, the maximum value of the standard deviation of any term increases as $r^{-r}\, g^{\frac{1}{2}}(r)$ which is seen from Table 4 to be rather slowly

TABLE 4.

Variance of Regression Coefficients from Simplex Centroid Design.

r	g(r)	$g^{\frac{1}{2}}(r)$	$r^{-r} g^{\frac{1}{2}}(r)$	$(1+e^{-2})^{r/2}$
1	1	1.00	1.00	1.07
2	24	4.90	1.22	1.14
3	1188	34.46	1.28	1.21
4	118400	344.09	1.34	1.29
5	19662000	4434.18	1.42	1.37

Plackett has pointed out that $r^{-r} g^{\frac{1}{2}}(r)$ can be approximated by $(1+e^{-2})^{r/2}$ and that for r = 7 agreement with the exact value is better than 2%. Figures calculated from this formula have been added to Table 4 for comparison.

We are now in a position to calculate the confidence limits on the regression coefficients or to ascertain if they are significantly different from zero by means of a t-test.

We have derived a regression equation for what was originally called a special cubic design but which is also the simplest simplex centroid design. The $100(1-2\alpha)\%$ confidence limits associated with the coefficients is $\beta \pm t\alpha.S.E(\beta)$ and the significance of the coefficients is measured by their ratio to their standard errors i.e. by

$$t = \beta/S.E.(\beta).$$

Now, the variance from Table 2 is 0.295 with 7 degrees of

freedom. $t_{0.025,7} = 2.365$. Hence the maximum 95% confidence limits for the coefficients are:

$$\beta_i \pm 2.365 \times 1.00 \times \sqrt{0.295} = \beta_i \pm 0.73 \qquad (1 \leq i \leq 3)$$

$$\beta_{ij} \pm 2.365 \times 1.22 \times \sqrt{0.295} = \beta_{ij} \pm 0.89 \qquad (1 \leq i < j \leq 3)$$

$$\beta_{123} \pm 2.365 \times 1.28 \times \sqrt{0.295} = \beta_{123} \pm 0.94$$

That is

β_1 : 100.9 ± 0.73, i.e. 100.2 to 101.6

β_2 : 85.4 ± 0.73, i.e. 84.7 to 86.1

β_3 : 85.5 ± 0.73, i.e. 84.8 to 86.2

β_{12} : -16.4 ± 0.89, i.e. -17.3 to -15.5

β_{13} : -10.8 ± 0.89, i.e. -11.7 to - 9.9

β_{23} : 0.0 ± 0.89, i.e. 0.9 to - 0.9

β_{123} : 29.2 ± 0.94, i.e. 28.3 to 30.2

The calculated t values are all highly significant except that of β_{23}, the value of which is zero. E.g. the t value of β_{13} is given by

$$t = -10.8/(1.22 \times \sqrt{0.295}) = -28.6$$

This is significant at the 0.00001% level! ($t_{0.0000001,7} = 19.932$).

It has been stressed that these figures are maximum confidence limits. If one wished to calculate the confidence limits on, say, $\beta_{123}(= 29.2)$ at the point (0.2, 0.3, 0.5) one would use the figures in the column $g^{\frac{1}{2}}(r)$ of Table 4 to obtain

$$29.2 \pm 2.365 \times (0.2 \times 0.3 \times 0.5) \times 34.46 \times \sqrt{0.295}$$

$$= 29.2 \pm 0.76 \quad \text{i.e. } 28.4 \text{ to } 2929.$$

Fractionation of Simplex Centroid Designs.

As the number of mixture variables (q) and process

variables (n) increases the number of points at which experimental observations have to be made increases rapidly being given by $(2^q-1)2^n$ when process variables are examined at only two levels.

If no process variables are present one may fractionate by dropping all terms higher than a certain chosen degree. For example the cubic regression equation is obtained by dropping all terms of degree > 3 to give

$$y = \sum_{1\leq i\leq q} \beta_i x_i + \sum_{1\leq i<j\leq q} \beta_{ij} x_i x_j + \sum_{1\leq i<j<k\leq q} \beta_{ijk} x_i x_j x_k$$

The usual precautions must be taken when fractionating; the terms being dropped must be known to be small enough to be ignored and one of the prices of economy is less precision.

Example.

The following data was obtained in an experiment carried out in developing a paint which would reduce decarburisation of steel during preheating.

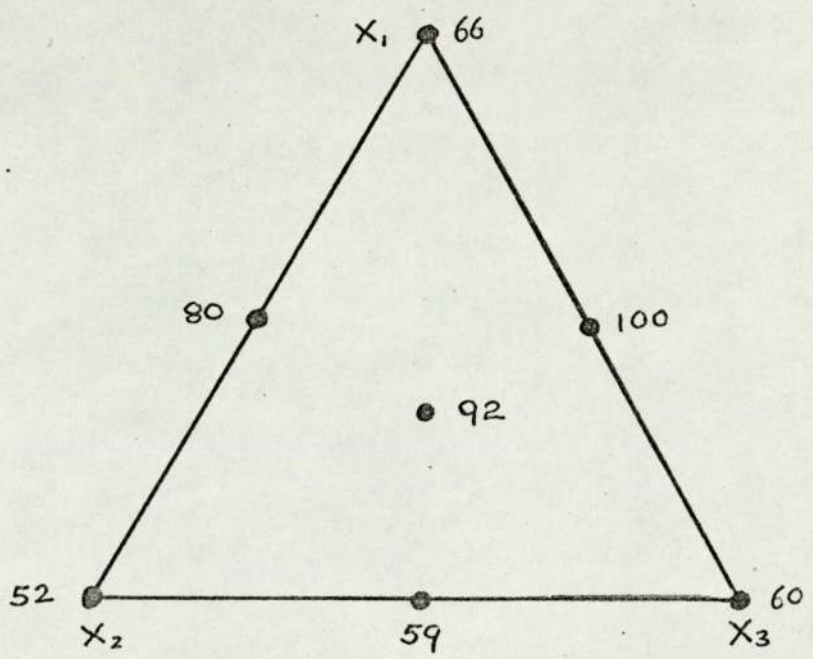


Fig 7: RESPONSE CONTOURS FOR UNFRACTIONATED DESIGN

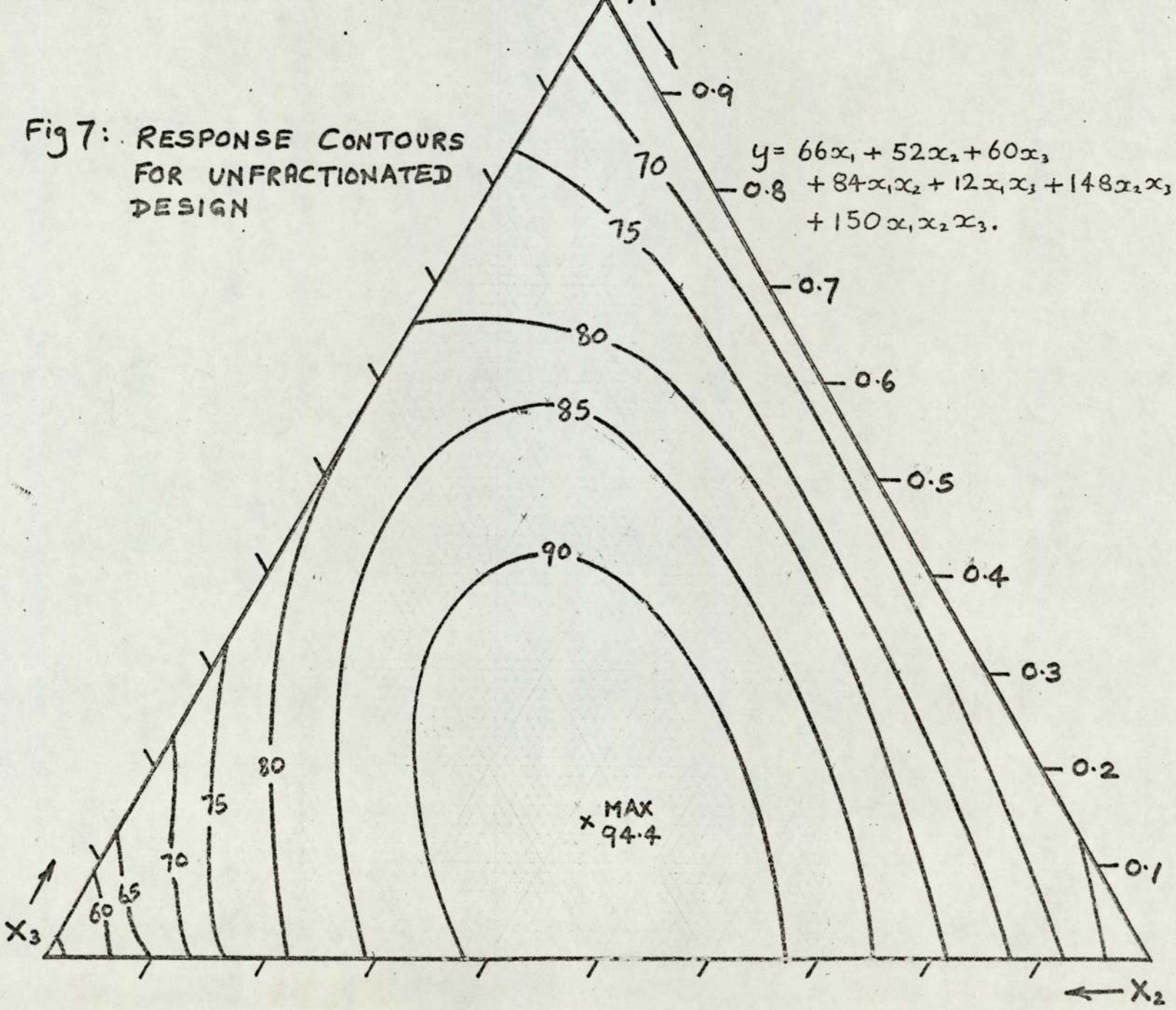
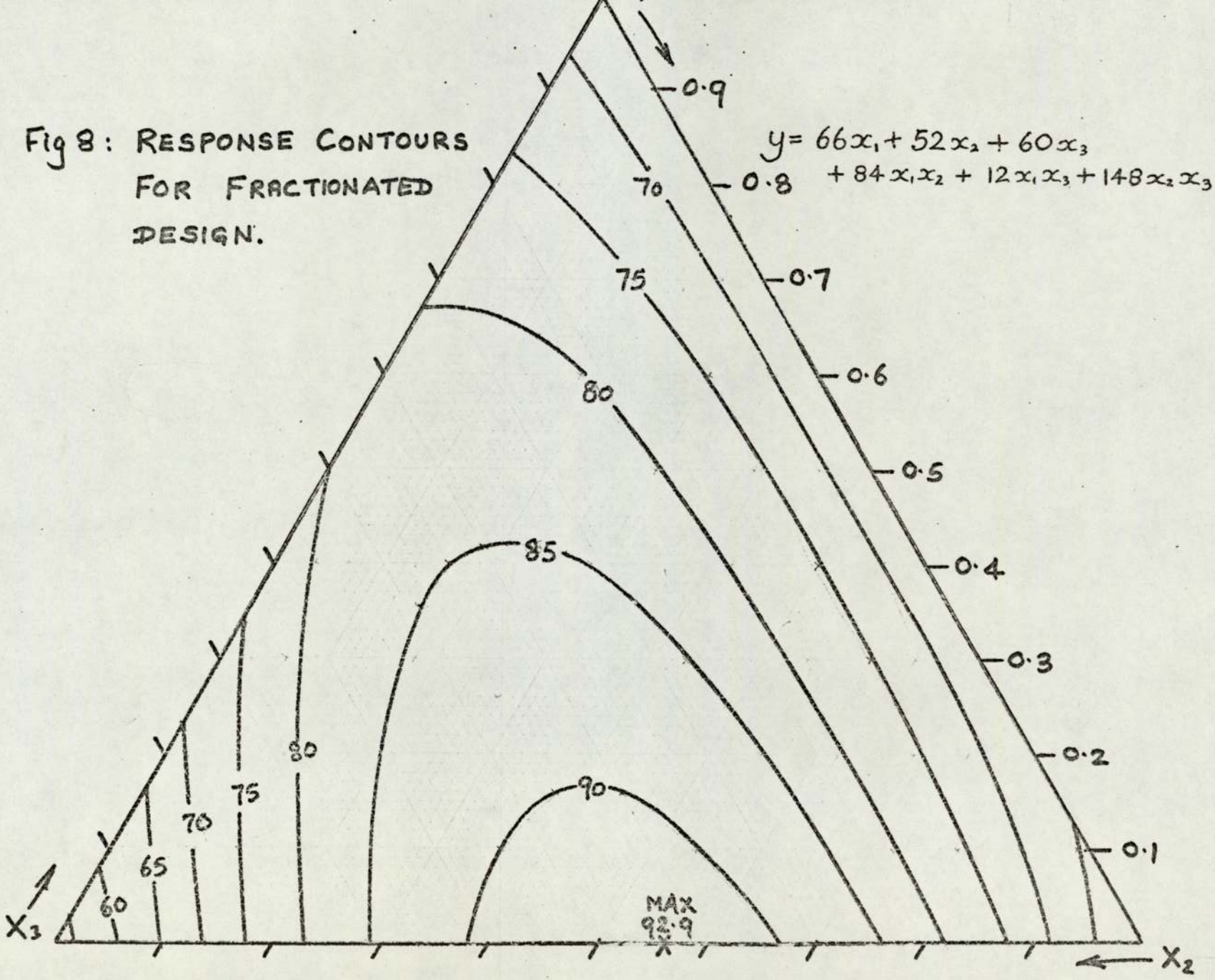


Fig 8: RESPONSE CONTOURS FOR FRACTIONATED DESIGN.

$$y = 66x_1 + 52x_2 + 60x_3 + 84x_1x_2 + 12x_1x_3 + 148x_2x_3$$



From the above the coefficients can be calculated

$\beta_1 = 66$ $\beta_{12} = 84$

$\beta_2 = 52$ $\beta_{13} = 12$ $\beta_{123} = 150$

$\beta_3 = 60$ $\beta_{23} = 148$

If a full cubic model is chosen the equation is

$$y_{ijk} = 66x_1 + 52x_2 + 60x_3 + 84x_1x_2 + 12x_1x_3 + 148x_2x_3 + 150x_1x_2x_3$$

If one decides to choose a quadratic model by dropping the last term the equation is

$$y_{ijk} = 66x_1 + 52x_2 + 60x_3 + 84x_1x_2 + 12x_1x_3 + 148x_2x_3$$

The response surface for these two equations is plotted in Figs. 7 and 8, and the misleading pattern produced by the fractionated equation can be appreciated. This is due to the fact that the coefficient associated with the cubic term is quite large. The important point here is that the data used in this example are actual experimental figures about which there was no previous knowledge. Had the design been fractionated in practice an erroneous conclusion would have been drawn.

When process variables are present a complication arises. In the first place when talking about the degree of the equation the process variables must be included. So that if there are q mixture variables and n process variables the degree of the equation is q+n. For instance in the example given on page 41.

$$\beta_{12}x_1x_2 = (-62.0+118.0A-150.0B+10.0AB)x_1x_2$$

and is of the fourth degree. So, if it is specified that an equation is to be of degree ≤ 3 one would have to eliminate the AB term from the binary coefficients and the A,B and AB terms from the ternary coefficients.

The second point is that if one is dealing with a

large number of process variables and one decides to reduce the number by fractionation it will be found that the fractional blocks at the various lattice points have certain combinations which are not in common. To illustrate this, suppose one wishes to examine 4 mixture variables and 5 process variables A,B,C,D,E at two levels and to keep in the regression equation all terms of degree ≤ 3. The variable coefficients will be

$$\beta_i = (I + A + B + C + D + E + AB + AC + AD + AE + BC + BD + BE + CD + CE + DE)$$

$$\beta_{ij} = (I + A + B + C + D + E)$$

$$\beta_{ijk} = I$$

Now, in a 2^5 experiment main effects and two factor interactions can be measured by a half replicate using the defining contrast ABCDE, assuming that three and four factor interactions are negligible. This is adopted for observations on the pure components. For the binary mixtures a quarter replicate will have to be chosen. The best is that with defining contrasts ABE, CDE, ABCD in which the main effects have 2-factor aliases. However, of one compares the principal blocks of the quarter and half replicates one finds four combinations in the former which are not given in the latter. This is shown in Table 5 in which the asterisks show the four in question. If one attempts to resolve the difficulty by generating a quarter replicate from the half replicate using the defining contrasts AB, CDE, ABCDE, the main effects A and B will be aliases of each other which makes the experiment pointless.

TABLE 5 : Principal Blocks of $\frac{1}{2}$ and $\frac{1}{4}$ replicates of a 2^5 experiment.

$\frac{1}{2} \times 2^5$				$\frac{1}{4} \times 2^5$	
(1)	ac	ae	ad	(1)	bce*
ab	bc	be	bd	ab	ade*
acde	de	cd	ce	cd	bde*
bcde	abde	abcd	abce	ace*	abcd

There are two possible ways to resolve this

(a) adopt the half replicate at both the binary and single component points, calculate the response and drop the unwanted terms,

(b) adopted the quarter replicate at the binary mixture points and add the four extra terms at the single component points to be used when necessary.

As (a) requires more observations than (b) and the object of fractionation is economy of effort and resources (b) is adopted.

At the ternary mixture point the problem is that of estimating the general mean of a 2^5 experiment and this can be done with a 2^{-5} fraction i.e. the single point (1) which is contained in all the other fractions.

In this example the full factorial would have taken $(2^4-1)2^5 = 480$ observations. With the above fraction we have

16+4	observations on	4 single components	= 80
8	"	" 6 binary mixtures	= 48
1	"	" 4 ternary mixtures	= 4
		Total:	132

An illustration of this ramification of the technique would have been useful at this point. But the amount

of computational work would be prohibitive within the time available.

One objection to these fractionated designs is the same as that against simplex lattice designs discussed on page 34 . If it is decided to keep the degree of the equation $\leq m$ then the prediction of the response is based on at most mixtures of m components. And the greater the difference between the number of components and the degree of the fractionated equation the less confidence one would have in the results. However, it may be a compromise between time and resources available and getting a bad answer to a problem which would otherwise have no answer at all.

Another objection based on intuitive grounds is the compromise one has to make with regard to the four treatments not in common in the quarter and half replicates. Scheffe gives two rules

1) As far as possible when fractions of process variables are of the same size they should be the same fraction.
2) As far as possible when fractions of process variables are of different sizes the larger fraction should contain the smaller.

We found that it was not sensibly possible to implement 2 as this gave a non-uniform distribution of points over the simplex which is intuitively unattractive.

Modifications of Simplex Lattice Designs.

In practice there are many situations where the whole of the simplex is not available to the experimenter. Certain mixtures may not be possible on technological grounds; they may be dangerous or unstable. On the other hand they may be uninteresting.

If, for instance, one were investigating the velocity of detonation of gunpowder, a three component mixture of sulphur, charcoal and potassium nitrate, one would not carry out any experiments on the pure components or binary mixtures because the velocity of detonation is zero.

Suppose that only the first component has to be restricted and is subject to the condition that $x_1 \geq h$. The second, third etc. components are replaced by "pseudocomponents" which are mixtures of the first component and proportions p_i of the other components

$$\sum_{i=2}^{q} p_i = 1-h, \qquad i > 1$$

This is shown in Fig. 7 where the restriction is that $x_1 \geq 40\%$

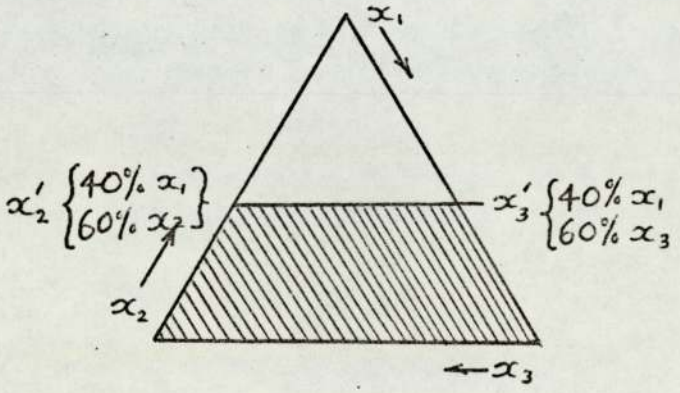


Fig 7

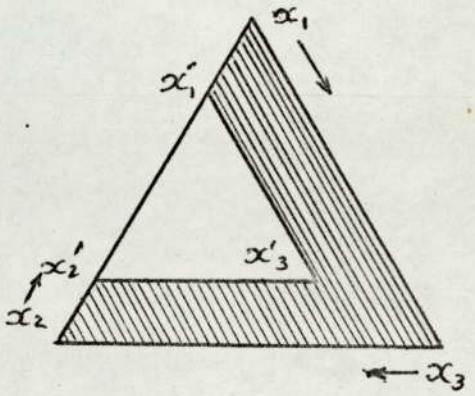


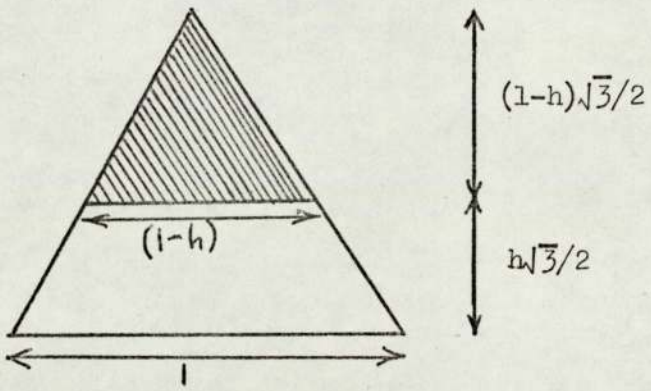
Fig 8

and the shaded area is not available. The experiments are carried out with the pure component x_1 and the two pseudo-components x_2' and x_3' and the necessary combinations demanded by the chosen regression equation. The idea can easily be extended to cover the situation shown in Fig.8 (See page 100 for a practical example of the use of this technique).

Consider now the alternative where $x_1 \leq h$. The

available factor space is the frustum of the simplex containing $h/[1-(1-h)^{q-1}]$ of the volume of the available factor space

Fig 9



This can easily be shown for the special case of a triangular simplex when $q = 3$. Consider the simplex of unit side in Fig. 9.

$$\text{Area top } \Delta = \tfrac{1}{2}(1-h)^2\,\sqrt{3}/2 = \sqrt{3}(1-h)^2/4$$

$$\text{Area whole } \Delta = \tfrac{1}{2}.1.\frac{\sqrt{3}}{2} = \frac{\sqrt{3}}{4}$$

$$\text{Area bottom trapezium} = \frac{\sqrt{3}}{4} - \frac{\sqrt{3}}{4}(1-h)^2$$

$$= 1-(1-h)^2\,\sqrt{3}/4$$

$$\text{Fractional area} = \frac{1-(1-h)^2\sqrt{3}/4}{\sqrt{3}/4} = 1-(1-h)^2$$

Now, the amount of x_1 is proportional to h and this has to be distributed within the fractional area available.

$$\frac{\text{Amount}}{\text{Fractional Area}} = \frac{h}{1-(1-h)^2}$$

and, in general $h/[1-(1-h)^{q-1}]$. For small h this can be expanded and second order and higher terms ignored to give $1/(q-1)$.

The solution to the problem of deriving a regression equation when $x_1 \leq h$ has been worked out only for the case of the quadratic polynomial (q,2) and a modified design exists which spreads the experimental points out into the corners of the frustum only if there is <u>one</u> small component (See Fig. 9).

Suppose $x_1 \leq h$; observations are taken of the responses of the pure components y_i $(i > 1)$ and the responses to their binary mixtures y_{ij} $(j > i > 1)$, the response y'_{ij} of the binary mixtures for which $x_1 = h$, $x_j = 1-h$ $(j > 1)$, and the responses of the y'_1 mixture with $x_1 = \frac{1}{2}h$ and $x_2 = x_3 = \ldots = x_q = (1-\frac{1}{2}h)/(q-1)$. Then the coefficients β_i $(i > 1)$ and β_{ij} $(j > i > 1)$ are still given by (9) and β_1 and β_{ij} are calculated from

$$h\beta_1 + h(1-h)\beta_{ij} = y'_{ij} - (1-h)\beta_j \qquad (j = 2,3, \ldots,q) \qquad (20)$$

and

$$\tfrac{1}{2}h\beta_1 + \frac{\frac{1}{2}h(1-\frac{1}{2}h)}{q-1}\sum_{2\leq j\leq q}\beta_{ij} = y'_1 - \frac{1-\frac{1}{2}h}{q-1}\sum_{2\leq j\leq q}\beta_j - \frac{(1-\frac{1}{2}h)^2}{(q-1)^2}\sum_{2\leq i\leq j\leq q}\beta_{ij} \qquad (24)$$

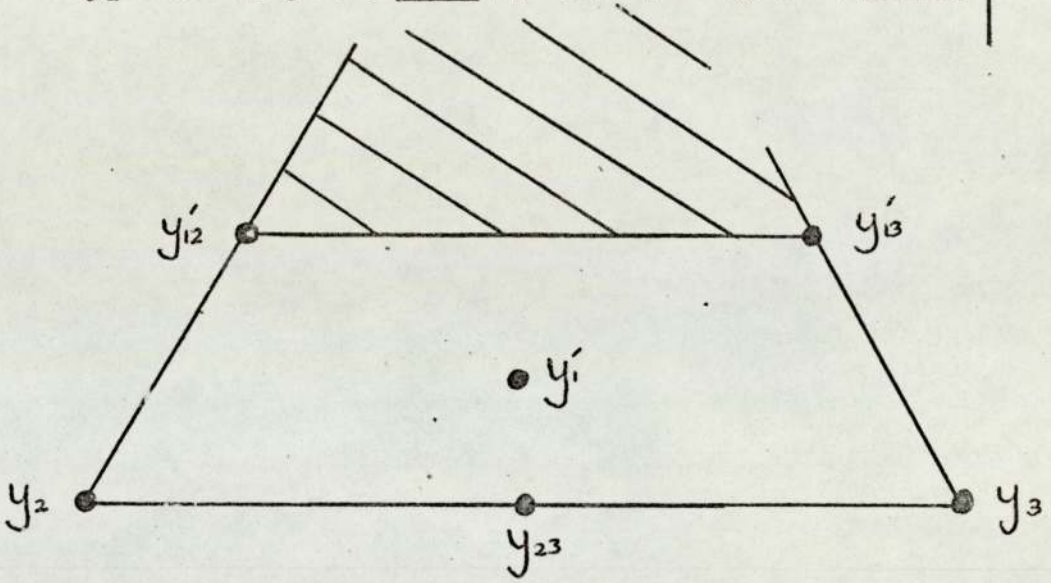
These may be solved for β_1 and β_{ij} to give the values of the regression coefficients.

Consider the case where $x_1 \leq 0.25$ on a (3,2) lattice i.e. $h = 0.25$.To examine this solution an equation was taken and the responses calculated for the points given above. Then, using (20) and (21), β_1,β_{12} and β_{13} were calculated and compared with the known values. The equation was

$$y = 90x_1 + 95x_2 + 100x_3 + 27x_1x_2 + 27x_1x_3 + 27x_2x_3$$

"Observation" were made at the following points i.e. the responses at the following points were calculated

	Responses
y_2 at $x_2 = \underline{1.0}$	95.0
y_3 at $x_3 = \underline{1.0}$	100.0
y_{23} at $x_2=x_3=\underline{0.5}$	104.25
y'_{13} at $x_1=0.25, x_2=1-0.25.\underline{0.75}$	102.5625
y'_{13} at $x_1=0.25, x_3=1-0.25=\underline{0.75}$	98.8125
y'_1 at $x_1=\frac{1}{2}\times 0.25=\underline{0.125}, x_2=x_3=(1-\frac{1}{2}\times 0.25)/(3-1)=\underline{0.4375}$	104.6836



Inserting these figures for the responses in (20) and (21) we obtain

(i) $\frac{1}{4}\beta_1 + \frac{1}{4}(1-\frac{1}{4})\beta_{12} = 98.8125 - (1-\frac{1}{4})\ 95$

i.e. $\underline{4\beta_1 + 3\beta_{12} = 441}$

(ii) $\frac{1}{4}\beta_1 + \frac{1}{4}(1-\frac{1}{4})\beta_{12} = 102.5625 - (1-\frac{1}{4})\ 100$

i.e. $\underline{4\beta_1 + 3\beta_{13} = 441}$

(iii) $\frac{1}{2}.\frac{1}{4}\beta_1 + \frac{\frac{1}{2}.\frac{1}{4}(1-\frac{1}{2}.\frac{1}{4})}{3-1}(\beta_{12}+\beta_{13}) = 104.6836 - \frac{(1-\frac{1}{2}.\frac{1}{4})}{3-1}(95+100) - \frac{(1-\frac{1}{2}.\frac{1}{4})^2}{(3-1)^2}\ 27$

i.e. $\underline{16\beta_1 + 7(\beta_{12}+\beta_{13}) = 1818}$

Solving this triplet of equations we obtain

$\beta_1 = 90$, $\beta_{12} = 27$ and $\beta_{13} = 27$

which are identical with the regression coefficients in the equation used to generate the observations, thus illustrating that it is not necessary to make an observation at $x_1 = 1$ in order to be able to estimate β_1.

q-component mixtures.

Lambrakis (1968a) has extended Scheffe's designs to the case where observations are made on mixtures containing all q components with non-zero proportions but has confined himself to simplex lattices of the (q,2) type i.e. quadratic models, because of the amount of algebra involved in deriving the necessary equations.

Letting the same letter represent the mixture and the response to the mixture the following are the new types of mixtures

y_i $(1 \leq i \leq q)$ is the response to a mixture with proportions

$$x_i = \tfrac{1}{2},\ x_r = \tfrac{1}{2}(q-1)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i;$$

y_{ij} $(1 \leq i < j \leq q)$ is the response to a mixture with proportions

$$x_i = x_j = \tfrac{1}{3},\ x_r = \tfrac{1}{3}(q-2)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j;$$

y_{iij} and y_{ijj} $(1 \leq i < j \leq q)$ are responses to mixtures with proportions

$$x_i = \tfrac{2}{4},\ x_j = \tfrac{1}{4},\ x_r = \tfrac{1}{4}(q-2)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j;$$

and $$x_i = \tfrac{1}{4},\ x_j = \tfrac{2}{4},\ x_r = \tfrac{1}{4}(q-2)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j;$$

y_{iij} and y_{ijjj} $(1 \leq i < j \leq q)$ are responses to mixtures with proportions

$$x_i = \tfrac{3}{5},\ x_j = \tfrac{1}{5},\ x_r = \tfrac{1}{5}(q-2)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j;$$

$$x_i = \tfrac{1}{5},\ x_j = \tfrac{3}{5},\ x_r = \tfrac{1}{5}(q-2)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j;$$

y_{ijk} $(1 \leq i < j < k \leq q)$ is the response to a mixture with proportions

$$x_i = x_j = x_k = \tfrac{1}{4},\ x_r = \tfrac{1}{4}(q-3)^{-1} \text{ for } 1 \leq r \leq q, r \neq i,j,k;$$

$y_{ijk\ell}(1 \leq i<j<k<\ell \leq q)$ in the response to a mixture with proportions

$$x_i=x_j=x_k=x_\ell = \tfrac{1}{5},\ x_r = \tfrac{1}{5}(p-4)^{-1} \text{ for } 1 \leq r \leq q,\ r \neq i,j,k,\ell;$$

and so on. So, for the case where q = 4, observations will be made at the following points

y_1 : $x_1 = \frac{1}{2}$, $x_2 = \frac{1}{6}$, $x_3 = \frac{1}{6}$, $x_4 = \frac{1}{6}$;

y_{12} : $x_1 = \frac{1}{3}$, $x_2 = \frac{1}{3}$, $x_3 = \frac{1}{6}$, $x_4 = \frac{1}{6}$;

y_{112} : $x_1 = \frac{1}{2}$, $x_2 = \frac{1}{4}$, $x_3 = \frac{1}{8}$, $x_4 = \frac{1}{8}$;

y_{122} : $x_1 = \frac{1}{4}$, $x_2 = \frac{1}{2}$, $x_3 = \frac{1}{8}$, $x_4 = \frac{1}{8}$;

y_{1112} : $x_1 = \frac{3}{5}$, $x_2 = \frac{1}{5}$, $x_3 = \frac{1}{10}$, $x_4 = \frac{1}{10}$;

y_{1222} : $x_1 = \frac{1}{5}$, $x_2 = \frac{3}{5}$, $x_3 = \frac{1}{10}$, $x_4 = \frac{1}{10}$;

y_{123} : $x_1 = \frac{1}{4}$, $x_2 = \frac{1}{4}$, $x_3 = \frac{1}{4}$, $x_4 = \frac{1}{4}$;

y_{1123} : $x_1 = \frac{2}{5}$, $x_2 = \frac{1}{5}$, $x_3 = \frac{1}{5}$, $x_4 = \frac{1}{5}$;

y_{1223} : $x_1 = \frac{1}{5}$, $x_2 = \frac{2}{5}$, $x_3 = \frac{1}{5}$, $x_4 = \frac{1}{5}$;

y_{1233} : $x_1 = \frac{1}{5}$, $x_2 = \frac{1}{5}$, $x_3 = \frac{2}{5}$, $x_4 = \frac{1}{5}$

and similarly for other combination of subscripts making a total of 39 observations, all on 4 component mixtures.* The position of the points lying in the x_1 plane are shown in Fig. 10.

For a quadratic model of a simplex lattice (9) is used which is given again

$$\hat{y} = \sum_{1 \leq i \leq q} \beta_i x_i + \sum_{1 \leq i<j \leq q} \beta_{ij} x_i x_j$$

The calculation of the coefficients is carried out as previously described by substituting the observed means and their proportions and after a great deal of algebra one obtains

---

*There are, in fact, 50 subscript combinations but y_{ijk}, y_{iijk}, and y_{ijjk} and y_{ijkk} all yield identical sets of values.

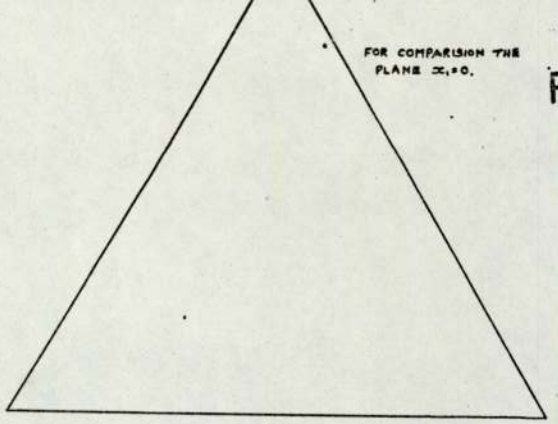
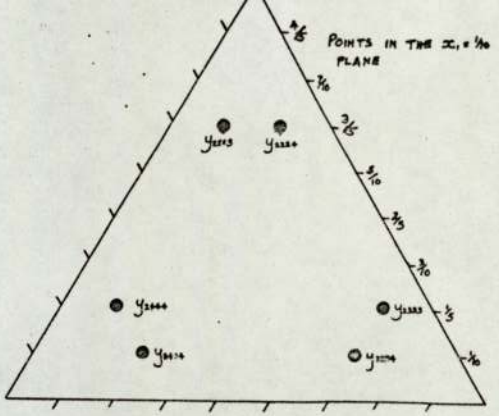
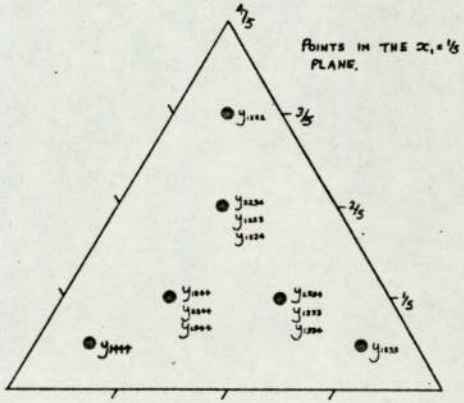
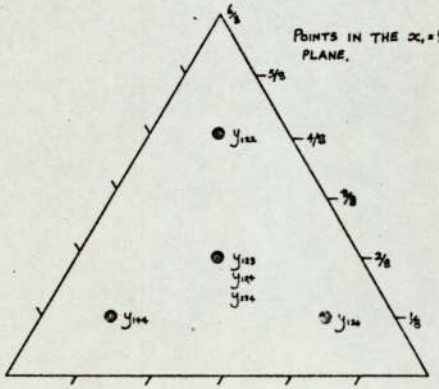
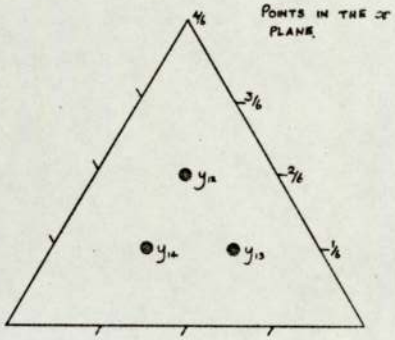

FOR COMPARISION THE PLANE $x_1 = 0$.

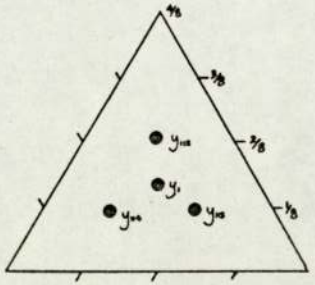
Fig. 10.

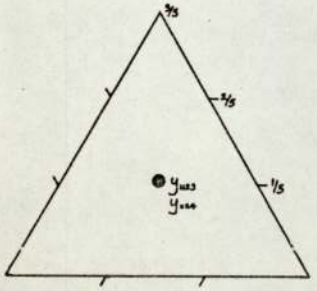

POINTS IN THE $x_1 = 1/10$ PLANE

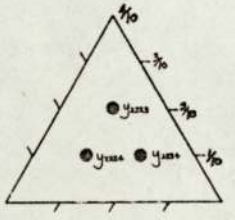

POINTS IN THE $x_1 = 1/5$ PLANE.


POINTS IN THE $x_1 = 1/4$ PLANE.


POINTS IN THE x PLANE.


POINTS IN THE $x_1 = 1/2$ PLANE.


POINTS IN THE $x_1 = 3/5$ PLANE.


POINTS IN THE x PLANE.

$$\beta_i = Ay_i + B \sum_{\substack{1\leq r<s\leq q \\ r,s\neq 1}} y_{rs} + C \sum_{i=1}^{q} y_i - D \sum_{1\leq i<j\leq q} y_{ij}$$

$$\beta_{ij} = -E(y_i+y_j)+Fy_{ij}+G \sum_{\substack{1\leq r<s\leq q \\ r,s\neq i,j}} y_{rs} - H\left(\sum_{\substack{1\leq r<s\leq q \\ r,s\neq j}} y_{rs} + \sum_{\substack{1\leq r<s\leq q \\ r,s\neq j}} y_{rs} \right)$$

$$- I \sum_{i=1}^{r} y_i + J \sum_{1\leq i<j\leq q}$$

where

$A = 4(q-1)^2(2q-5)/(q-2)(q^2-2q-2)$

$B = 9(q-2)^2/(q-3)(q^2-2q-2)$

$C = 4(q-1)(5q^4-55q^3+227q^2-417q+288)/(q-2)(q-3)(q^2-2q-2)$
$(q^2+2q-12)$

$D = 9(q^5+q^4-58q^3+240q^2-372q+200)/(q-1)(q-3)(q^2-2q-2)(q^2+2q-12)$

$E = 12(q-1)^2/(q^2-2q-2)$

$F = 18(q-2)^2/(q-3)(2q-5)$

$G = 9(q-2)^2/(q-3)^2(2q-5)$

$H = 27(q-2)^3/(q-3)(2q-5)(q^2-2q-2)$

$I = 24(q-1)(4q^4-48q^3+213q^2-415q+300)/(q-3)(2q-5)(q^2-2q-2)$
$(q^2+2q-12)$

$J = 9(q-2)(5q^5+2q^4-298q^3+1348q^2-2240q+1264)/(q-1)$
$(q-3)(2q-5)(q^2-2q-2)(q^2+2q-12)$

and r,s = i means that one of r and s is equal to i, and r,s = i,j means that one of r and s is equal to i or j. In the same way equations can be derived for the variance etc. and these are fully and clearly given in Lambrakis's paper for the (q,2) case.

However, the equations apply only for mixtures of 4 components or more for, if q = 3, all the coefficients given above, apart from A and E, will be infinite due to the term (q-3)

in the denominator.

While the rationale of this approach can be appreciated - that if one is investigating mixtures of q components then measurements should be made only on mixtures of q components - it leads to some tedious coefficients and more complex expressions than either the simplex lattice or the simplex centroid designs. Further, the number of observation points is large. For this design 4 components would require 39 observations; a simplex centroid design would require $2^4-1 = 15$ observations while a quartic polynomial could be fitted using a simplex lattice (see Table 1).

Table 6 gives values of the coefficients A to J for q = 4(1)10.

TABLE 6 : Lambrakis's Coefficients.

	q = No. of components						
Coefficient.	4	5	6	7	8	9	10
A	9.0	8.2051	7.9545	7.8545	7.8116	7.7939	7.7885
B	6.0	3.1154	2.1818	1.7045	1.4087	1.2049	1.0549
C	-4.0	0.6065	1.0480	1.1209	1.1159	1.0871	1.0499
D	5.0	3.1605	2.3879	1.9318	1.6237	1.3998	1.2291
E	2.0	0.9231	0.5455	0.3636	0.2609	0.1967	0.1539
F	24.0	16.200	13.714	12.500	11.782	11.308	10.971
G	12.0	4.0500	2.2857	1.5625	1.1782	0.9423	0.7837
H	12.0	5.6077	3.7403	2.8409	2.3051	1.9464	1.6879
I	0.0	1.6054	2.1212	2.3102	2.3632	2.3516	2.3077
J	12.0	1.6508	5.8251	4.7360	3.9940	3.4513	3.0359

Double Lattices.

Lambrakis (1968b) developed a theory for experiments with mixtures where each component itself is a mixture of several other components. An example might be a two polymers blends, the first composed of polymer A, plasticiser B and filler B and the second composed of polymer W, polymer X, plasticiser Y and

filler Z and it is desired to obtain data on the physical properties of mixtures of these two polymer blends.

If we have one simplex lattice x_1, x_2, x_3 for the first set of components and other simplex lattice z_1, z_2, z_3, z_4 for the second set of components and also all possible mixtures which can be produced by mixing each mixture from the first simplex lattice with each mixture from the second simplex lattice with proportions c_1 and c_2 respectively ($c_1+c_2 = 1$) then we shall have a double lattice and if the first lattice was for a quadratic model (q,2) and the second for a cubic model (p,3) we would have a (p,q;3,2) double lattice which, in the special case above where $q = 3$ and $p = 4$, would be a (3,4;3,2) double lattice. If we were dealing with a simplex lattice design this would mean that observations would have to be made at $6 \times 20 = 120$ points (see Table 2), while if a simplex centroid design were used there would be $(2^3-1)(2^4-1) = 105$ points. Taking into account the replication necessary to give a measure of variance it is obvious that we have a very unweildy experiment and there is a strong case for fractionation.

The double lattice polynomials are obtained by multiplying together the separate polynomials for each lattice.

The polynomial for the (q,2) simplex lattice is given by

$$y = \sum_{1 \leq i \leq q} \beta_i z_i + \sum_{1 \leq i < j \leq q} \beta_{ij} z_i z_j$$

and that for the (q,3) simplex lattice by

$$y = \sum_{1 \leq i \leq q} \beta_i x_i + \sum_{1 \leq i < j \leq q} \beta_{ij} x_i x_j + \sum_{1 \leq i < j \leq q} y_{ij} x_i x_j (x_i - x_j) + \sum_{1 \leq i < j < k \leq q} \beta_{ijk} x_i x_j x_k$$

Multiplying the two polynomials together and replacing the products

of the coefficients by a single coefficient, we obtain the regression equation for the (p,q;3,2) double lattice

$$y = \sum_{1\leq i\leq p}\sum_{1\leq j\leq q}\beta_{i,k}x_i z_j + \sum_{1\leq i\leq p}\sum_{1\leq j<k\leq q}\beta_{i,jk}x_i z_j z_k$$

$$+ \sum_{1\leq i<j\leq p}\sum_{1\leq k\leq q}\beta_{ij,k}x_i x_j x_k + \sum_{1\leq i<j\leq p}\sum_{1\leq k<\ell\leq q}\beta_{ij,k\ell}x_i x_j z_k z_\ell$$

$$+ \sum_{1\leq i<j<p}\sum_{1\leq k\leq q}\gamma_{ij,k}x_i x_j (x_i - x_j) z_k + \sum_{1\leq i<j\leq p}\sum_{1\leq k<\ell\leq q}\gamma_{ij,k\ell}x_i x_j (x_i - x_j) z_k z_\ell$$

$$+ \sum_{1\leq i<j<k\leq p}\sum_{1\leq \ell\leq q}\beta_{ijk,\ell}x_i x_j x_k x_\ell$$

$$+ \sum_{1\leq i<j<k\leq p}\sum_{1\leq \ell<m\leq q}\beta_{ijk,\ell m}x_i x_j x_k z_\ell z_m$$

This formula has been quoted to illustrate the complexity of the formulae involved with a relatively simple case of a multiple lattice. Further, the equations for estimating the coefficients, apart from the first four, are arithmetically tedious, the one for $\gamma_{ij,k\ell}$ having 12 terms, that for $\beta_{ijk,\ell}$ having 10 terms and that for $\beta_{ijk,\ell m}$ having 30 terms.

While this work is didactically valuable in that it can be considered as the general case of the simplex lattice design, the writer feels that in practice it would be computationally simpler to work out the various mixtures of the pq single components and use these as pseudocomponents in a

normal simplex lattice (pq,m) or a simplex centroid design. Strictly speaking this does not answer the original question of the effect of mixing two mixtures. But it is computationally more attractive and would probably yield a practical meaningful result though not satisfying the purist.

Bias and Variance-free Designs.

Draper and Lawrence (1965a) show that Scheffe's designs are "all variance" designs and are not suitable for situations where both variance and bias exist. In their 1965a paper they derive suitable designs for three components and in their 1965b paper, four components.

These two authors find it easier to work in cartesian coordinates with the origin at the centroid of the simplex whose coordinates are

$$(0, m/\sqrt{3}) \quad , \quad (m/2, -m\sqrt{3}/6) \quad , \quad (-m/2, -m\sqrt{3}/6)$$

where m is the length of the side of the equilateral triangle, which in this work was 1 decimetre.

Box and Draper (1959) suppose the response surface is a polynomial of degree d_1 in x

$$\hat{y}(\underline{x}) = \underline{x}_1' \, \underline{\beta}_1$$

while the true function over the region of interest, in this case a triangular simplex, is a polynomial of degree d_2:

$$\eta(\underline{x}) = \underline{x}_1' \, \underline{\beta}_1 + \underline{x}_2' \, \underline{\beta}_2$$

We wish to arrange for the difference $\hat{y}(\underline{x}) - \eta(\underline{x})$ to be as small as possible and the measure of "closeness" used is

$$E[\hat{y}(\underline{x}) - \eta(\underline{x})]^2$$

Over the whole region of interest this is averaged out to give

$$J = \int_R E[\hat{y}(\underline{x}) - \eta(\underline{x})]^2 \Big/ \int_R dx$$

where R is the region of interest. In this case the region of interest is an equilteral triangle and $\int_R = \int_{-a}^{2a} \int_{-b}^{b}$.

Now, the average variance is given by

$$V = \int_R \left[\hat{y}(\underline{x}) - E\, \hat{y}(\underline{x}) \right]^2 dx \Big/ \int_R dx$$

and the average squared bias is given by

$$B = \int_R \left[E\, \hat{y}(\underline{x}) - \eta(\underline{x}) \right]^2 dx \Big/ \int_R dx$$

and it can be shown that J= V+B and that to minimise the bias one must minimise B.

Proceeding along these lines, Draper and Lawrence (1965a,b) calculate five point sets which define the coordinates of the observation points.

Set 1 : vertices of an equilateral triangle, centroid at the origin, side p:

$$\left(0, \frac{1}{\sqrt{3}} p\right), \left(\pm \tfrac{1}{2}p, -\frac{\sqrt{3}}{6} p\right)$$

Set 2 : vertices of an equilateral triangle, inverted with respect to Set 1, centroid at the origin, side q

$$\left(0, \frac{1}{\sqrt{3}} q\right), \left(\pm \tfrac{1}{2}q, \frac{\sqrt{3}}{6} q\right).$$

Set 3 : vertices of a square, side a, centroid at the origin, sides parallel to coordinate axes: $(\pm\, á, \pm\, a)$

Set 4 : points on the coordinate axes at a distance b from the origin: $(\pm b,0)$, $(0,\pm b)$.

Set 5 : vertices of a rectangle: $(c,d),(-d,c),(-d,-d),(d,-c)$

p,q,a,b,c and d are tabulated for certain selected values for two cases

a) when the model is linear ($d_1=1$) and there is a possible error due to the presence of second degree terms ($d_2=2$) in the true model and

b) the model is quadratic ($d_1=2$) and there is a possibility of error due to a third degree term ($d_2=3$).

For case (a) figures are given for designs incorporated 5-9 points (solutions for value < 5 do not exist) and for case (b) approximate designs are given for 7-12 points and accurate designs for 13-15 points.

Example $d_1=1$, $d_2=2$

Point sets: 1 and 2

p = 0.662, q = 0.381

No. of observations: 7

Coordinates: $\left(0, \frac{1}{\sqrt{3}} \times 0.662m\right), \left(\pm\frac{1}{2}\ 0.662m, -\frac{\sqrt{3}}{6} \times 0.662m\right)$

$\left(0, \frac{1}{\sqrt{3}} \times 0.381m\right), \left(\pm\frac{1}{2} \times 0.381m, \frac{\sqrt{3}}{6} \times 0.381m\right)$

i.e. (0,0.382m), ($\pm$ 0.331m, -0.191m)

(0,0.220m), ($\pm$ 0.195, 0.110m) and centroid

Example $d_1=1$, $d_2 = 2$

Point sets: 1, 1 and 2.

No. of observations: 9

$p_1 = 0.606m$, $p_2 = 0.550m$, $q = 0.364m$

Coordinates: (0,0.347m), ($\pm$0.303m, -0.175m)

(0,0.289m), ($\pm$0.250m, -0.144m)

(0,0.210m), ($\pm$0.182m, 0.105m)

Example $d_1=2$, $d_2=3$

Point sets: 1,3 and 4

No. of observations: 13

$p = 0.756m$, $a = 0.183m$, $b = 0.258m$

Coordinates (0,0.436m), ($\pm$0.378m, -0.218m)

($\pm$ 0.183m, $\pm$0.183m)

($\pm$ 0.258m,0), (0, $\pm$0.258m) and two points at centroid

Example $d_1=2$, $d_2=3$

Point sets: 1,1,2,5

$p_1 = 0.545m$, $p_2 = 0.776m$, $q = 0.480m$

$c = 0.071m$, $d = 0.171m$

No. of observations = 15

Coordinates: (0,0.314m), ($\pm$0.272m, -0.157m),

(0,0.448m), ($\pm$0.388m, -0.224m),

(0,-0.277m),($\pm$0.240m, 0.139m),

(0.071m,0.171m),(-0.171m, 0.071m)

(-0.071m,-0.171m),(0.171m,-0.071m) and two points at centroid

These coordinates are illustrated in Figs. in which m = 20 cm. and the figures calculated above are adjusted accordingly.

When both bias and variance errors are present each design point (x_1,x_2) is replaced by $(\theta x_1, \theta x_2)$ where $\theta > 1$

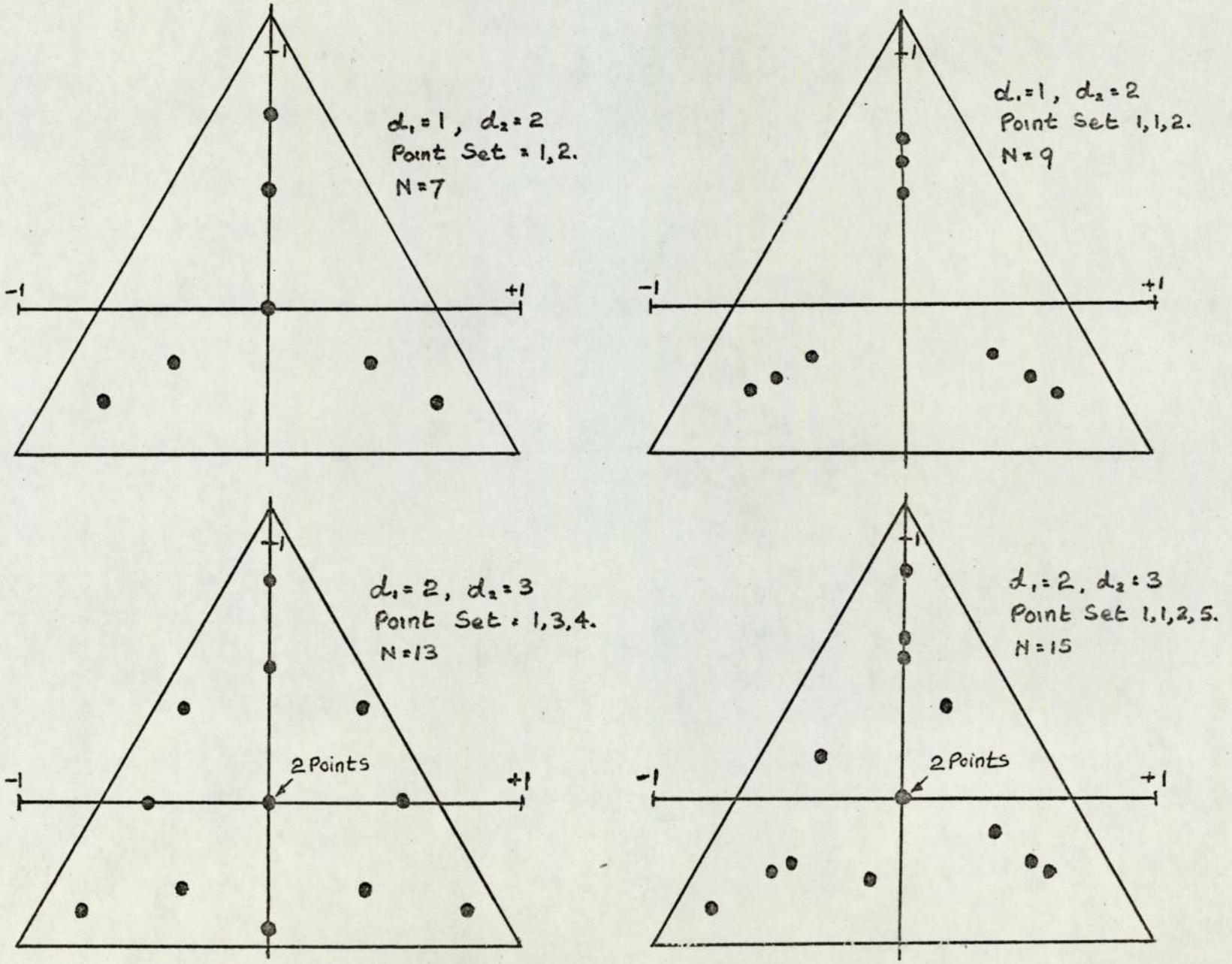

$d_1 = 1$, $d_2 = 2$
Point Set = 1, 2.
N = 7
$d_1 = 1$, $d_2 = 2$
Point Set 1, 1, 2.
N = 9
$d_1 = 2$, $d_2 = 3$
Point Set = 1, 3, 4.
N = 13
2 Points
$d_1 = 2$, $d_2 = 3$
Point Set 1, 1, 2, 5.
N = 15
2 Points

Fig 11

is a scale factor. In all-bias designs $\theta = 1$ and in all-variance designs $\theta = \infty$. Draper and Lawrence give tables from which it is possible to calculate θ in the cases $d_1=1$, $d_2=2$ and $d_1=2$, $d_2=3$. But these require some a priori knowledge of the β coefficients in the polynomial. In the absence of such knowledge they suggest that in the first case, as a rough rule of thumb $\theta = 1.1$ and in the second case $\theta = 1.2$. If, in a mixture problem a design cannot be expanded without some of the points going outside the region, a slightly smaller value of θ can be used. Identical figures are used for the 4 component case.

A comparison of Draper's and Scheffe's designs.

To compare Draper's designs with Scheffe's as approximating functions for a cubic polynomial, the cubic polynomial given by Gorman and Hinman (1962) and illustrated in Fig. 12 was taken. The quadratic approximation is obtained by ignoring all third order terms giving as the response function

$$y = 54.91x_1 + 3.89x_2 + 9.87x_3$$
$$-44.56x_1x_2 - 28.70x_1x_3 + 21.49x_2x_3$$

and is illustrated in Fig.13.

To illustrate Draper's method the 13 point set of Fig.11 was chosen. The cartesian coordinates were first converted to triangular coordinates and the relevant responses at these points were obtained using the equation of Fig. 12 and are given in Table 7.

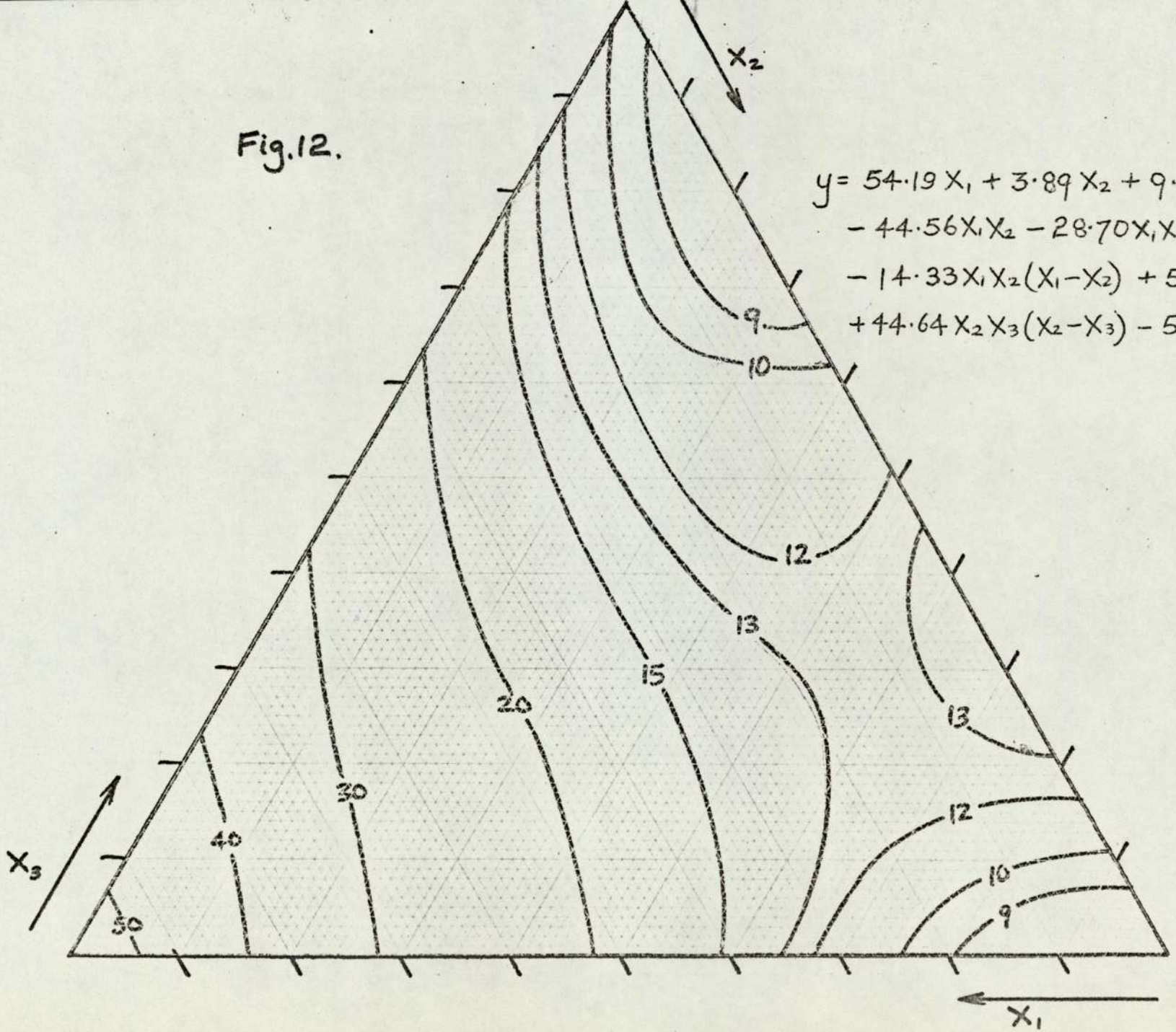
The regression coefficients were obtained using the formula $\underline{\beta} = (\underline{X}'\underline{X})^{-1}\underline{X}'\underline{Y}$ and the analysis of variance table obtained in the usual way.

y	X_0	X_1	X_2	X_1^2	X_2^2	X_1X_2
12.1	1	0.258	0.0	0.0666	0.0	0.0
27.77	1	-0.258	0.0	0.0666	0.0	0.0
10.94	1	0.0	0.258	0.0	0.0666	0.0
18.14	1	0.0	-0.258	0.0	0.0666	0.0
10.86	1	0.183	0.183	0.0335	0.0335	0.0335
12.79	1	0.183	-0.183	0.0335	0.0335	-0.0335
20.04	1	-0.183	0.183	0.0335	0.0335	-0.0335
26.21	1	-0.183	-0.183	0.0335	0.0335	0.0335
15.08	1	0.0	0.0	0.0	0.0	0.0
15.08	1	0.0	0.0	0.0	0.0	0.0
9.39	1	0.0	0.436	0.0	0.190	0.0
41.46	1	-0.378	-0.218	0.143	0.0475	0.0824
9.50	1	0.378	-0.218	0.143	0.0475	-0.0824

Table 6

TABLE 7. OBSERVATIONS AND CO-ORDINATES FOR THIRTEEN POINT DESIGN.

Fig.12.



$$y = 54{\cdot}19X_1 + 3{\cdot}89X_2 + 9{\cdot}87X_3 - 44{\cdot}56X_1X_2 - 28{\cdot}70X_1X_3 + 21{\cdot}49X_2X_3 - 14{\cdot}33X_1X_2(X_1-X_2) + 5{\cdot}13X_1X_3(X_1-X_3) + 44{\cdot}64X_2X_3(X_2-X_3) - 55{\cdot}54X_1X_2X_3.$$

Fig.13.

$$y = 54{\cdot}91X_1 + 3{\cdot}98X_2 + 9{\cdot}87X_3 - 44{\cdot}56X_1X_2 - 28{\cdot}70X_1X_3 + 21{\cdot}49X_2X_3.$$

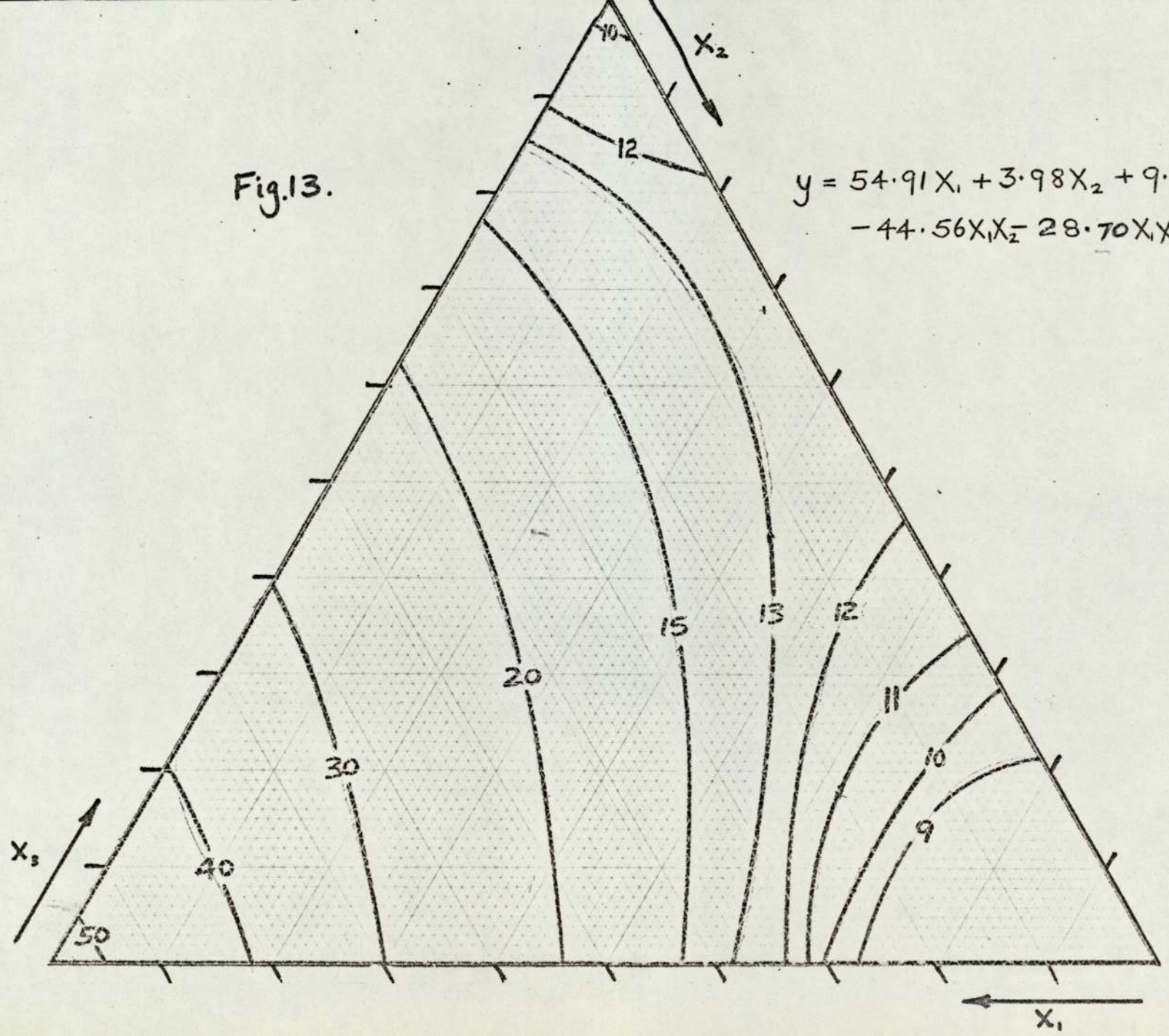
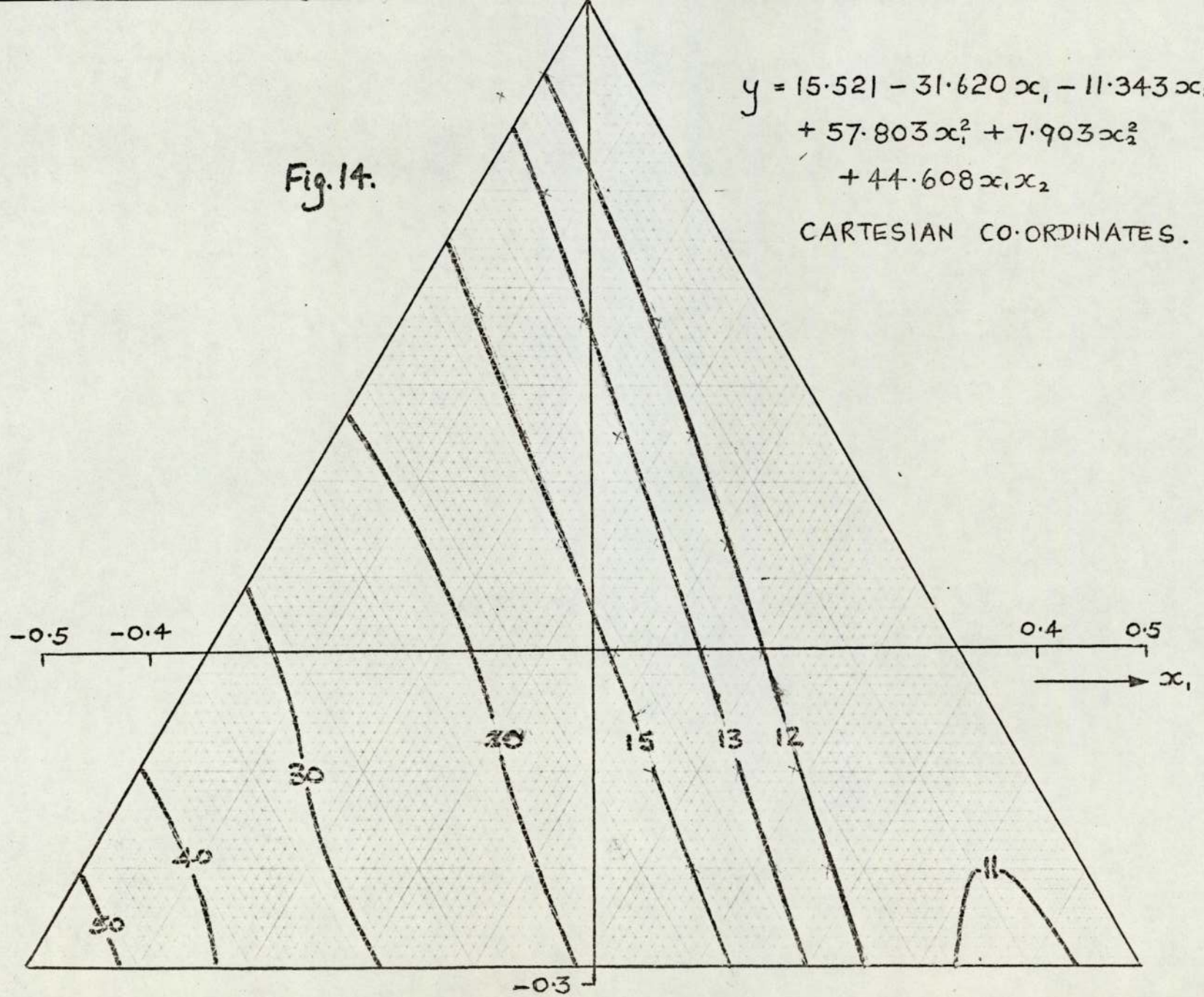


Fig. 14.

$$y = 15{\cdot}521 - 31{\cdot}620\,x_1 - 11{\cdot}343\,x_2 + 57{\cdot}803\,x_1^2 + 7{\cdot}903\,x_2^2 + 44{\cdot}608\,x_1x_2$$

CARTESIAN CO-ORDINATES.



Analysis of Variance

SOURCE	S.S.	D.F.	M.S.	F
β_0	4045.20	1		
$\beta_1,\beta_2 \mid \beta_0$	938.01	2	469.01	744.5
$\beta_{11},\beta_{22},\beta_{12} \mid \beta_1,\beta_2,\beta_0$	100.98	3	33.66	53.4
Residual	4.39	7	0.63	
	5088.58			

The response function is

$$y = 15.521 - 31.620x_1 - 11.343x_2$$
$$+57.803x_1^2 + 7.903x_2^2 + 44.608x_1x_2$$

where the x's in this case are measured on the cartesian system. This is illustrated in Fig. 14.

Of the variation about the mean R^2=938.01/(5088.58-4045.20) = 0.899 (or 89.9%) is explained, a very high proportion.

From the equation of Fig. 11 the expected values of the responses at the coordinates of Table 7 can be obtained. These are given in Table 8 together with the observed values obtained using Draper's equation and Scheffe's equation

Table 8.

Expected	Observed (Scheffe)	Observed (Draper)
12.10	11.26	11.21
27.77	28.95	27.53
10.94	14.72	13.12
18.14	18.22	18.97
10.86	12.75	11.35
12.79	11.32	12.51
20.04	21.38	19.94
26.21	27.75	27.08
15.08	17.14	15.52
15.08	17.14	15.52
9.39	12.27	12.08
41.46	42.24	42.26
9.50	6.77	11.00

The values of χ^2 for these two sets of observation are

$$\left.\begin{aligned} \chi^2(\text{Scheffe}) &= 1.80 \\ \chi^2(\text{Draper}) &= 4.34 \end{aligned}\right\} \text{ with 12 d. of f.}$$

Neither is significant ($\chi^2_{12,0.05} = 5.23$), but the figure for Draper's design is better than that for Scheffe's.

If, from the calculation of χ^2, one picks out the three points which make the largest contribution one finds they are the same for both sets of data, but not in the same numerical order and they correspond to areas on the top and bottom right of the simplex, the more complex part of the response surface. This is where their greatest deficiency lies.

The two equations can be qualititively compared by converting, say, Scheffe's triangular coordinates to cartesian coordinates using the relationship

$$z_1 = (-3x_1 - x_2\sqrt{3}+1)/3,\ z_2 = (3x_1 - x_2\sqrt{3}+1)/3,\ z_3 = (2x_2\sqrt{3}+1)/3$$

where z_1, z_2, z_3 are triangular coordinates and x_1, x_2 are cartesian coordinates. If this is done we obtain

$$y = 17.17 - 34.19x_1 - 6.82x_2 + 44.56x_1^2 - 10.05x_2^2 + 57.94x_1x_2$$

as compared with Draper's

$$y = 15.52 - 31.62x_1 - 11.34x_2 + 27.8x_1^2 + 7.903x_2^2 + 44.61x_1x_2$$

The most outstanding difference is in the coefficient of x_2^2. In Scheffe's equation it is almost as antagonistic as it is synergistic in Draper's.

Bias in Simplex Designs.

Bax and Draper (1959) have pointed out that in practice a graduating function such as a polynomial will always fail to some extent to represent a true function (see page 68). This they call "bias error".

To examine this a quartic equation given by Gorman and

Heinman (1962) was used as a "true function". This is shown in Fig. 7. The equation was evaluated for a series of points over the simplex and from the computer print-out the calculated responses at the points necessary for the calculation of the coefficients of a linear, quadratic, special cubic and cubic polynomial were read off and the relevant coefficients calculated. These are given in Table 9.

TABLE 9.

Responses

$y_1 = 95.03$	$y_{12} = 94.50$	$y_{112} = 94.35$	$y_{133} = 98.42$
$y_2 = 94.97$	$y_{13} = 99.67$	$y_{122} = 94.46$	$y_{223} = 95.14$
$y_3 = 94.08$	$y_{23} = 96.37$	$y_{113} = 98.29$	$y_{233} = 96.35$

$y_{123} = 95.49$

Coefficients

(i) Linear : $\beta_1 = 95.03$ $\beta_2 = 94.97$ $\beta_3 = 94.08$

(ii) Quadratic : $\beta_{12} = -2.00$ $\beta_{13} = 20.46$ $\beta_{33} = 7.38$

(iii) Sp.Cubic : $\beta_{123} = -56.01$

(iv) Cubic : $\beta_{12} = 2.668$ $\beta_{13} = 17.10$ $\beta_{23} = 5.468$

$\gamma_{12} = -0.8775$ $\gamma_{13} = -3.420$ $\gamma_{33} = -10.10$

$\beta_{123} = 38.16$

The response surfaces obtained using these coefficients are given in Fig. 15, 16 and 18 . The difference amongst the response surfaces is, in the writer's opinion, startling. So much so that the computer programs we carefully checked and the coefficients re-calculated by an independent worker. The "hill" persists in all of them being a maximum at about $x_1 = 0.5$. But the long valley on the left with a minimum near its head disappears in the cubic response surface to be replaced by a much shorter wider valley. The special cubic represents this almost as a basin while the quadratic causes it to degenerate to a long shallow slope.

Fig. 15

QUARTIC RESPONSE SURFACE

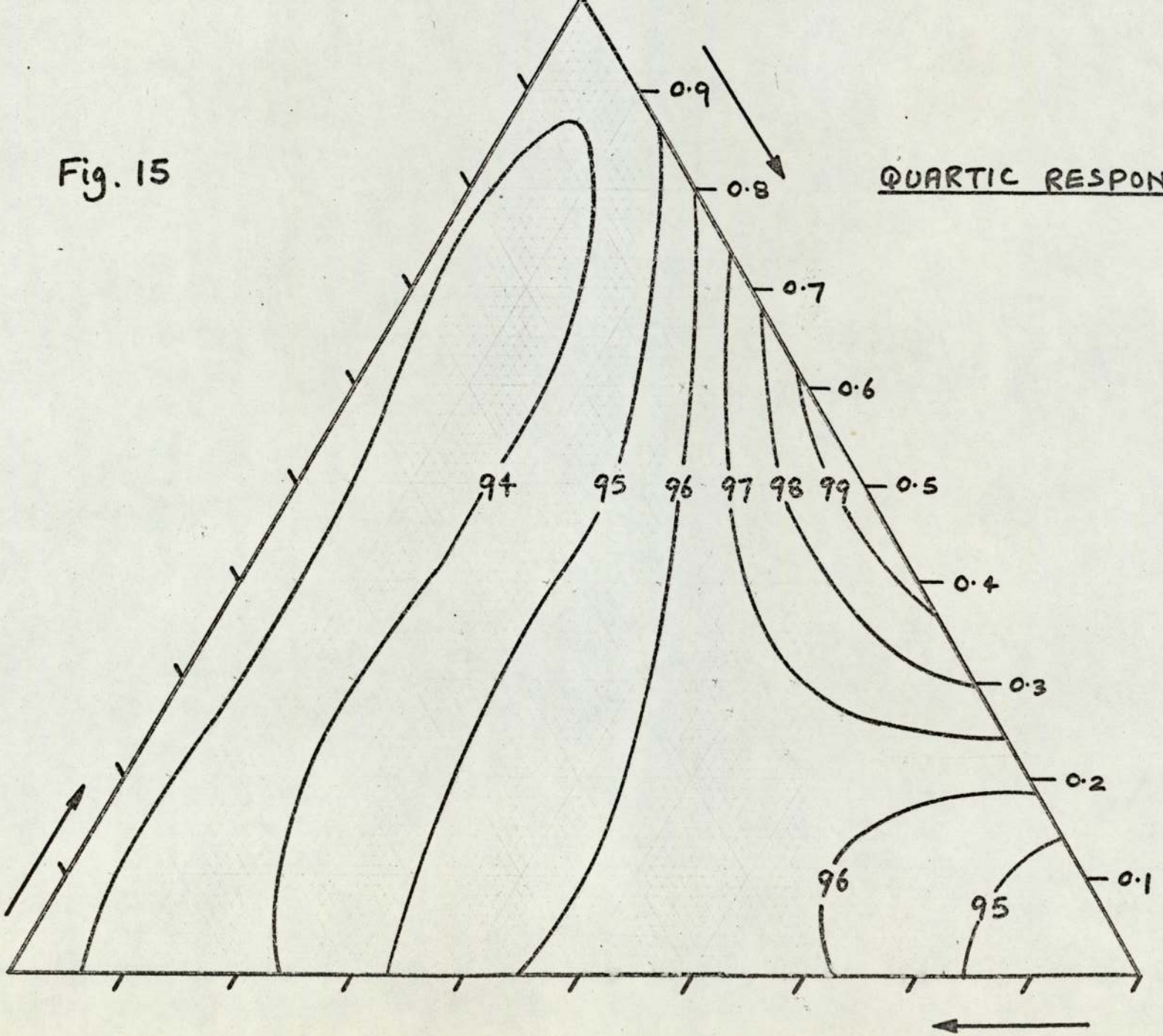
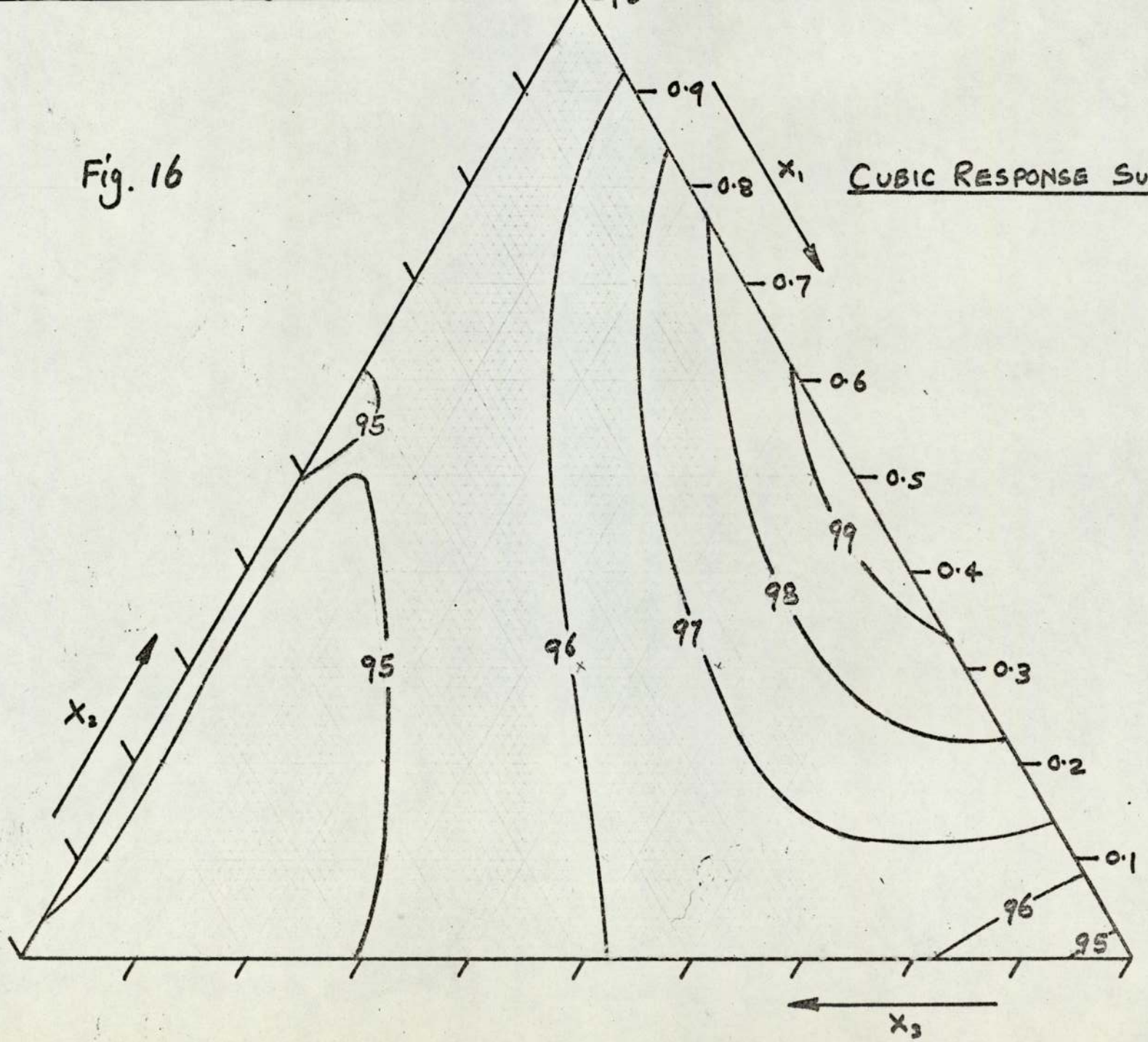


Fig. 16



CUBIC RESPONSE SURFACE.

Fig. 17

QUADRATIC RESPONSE SURFACE

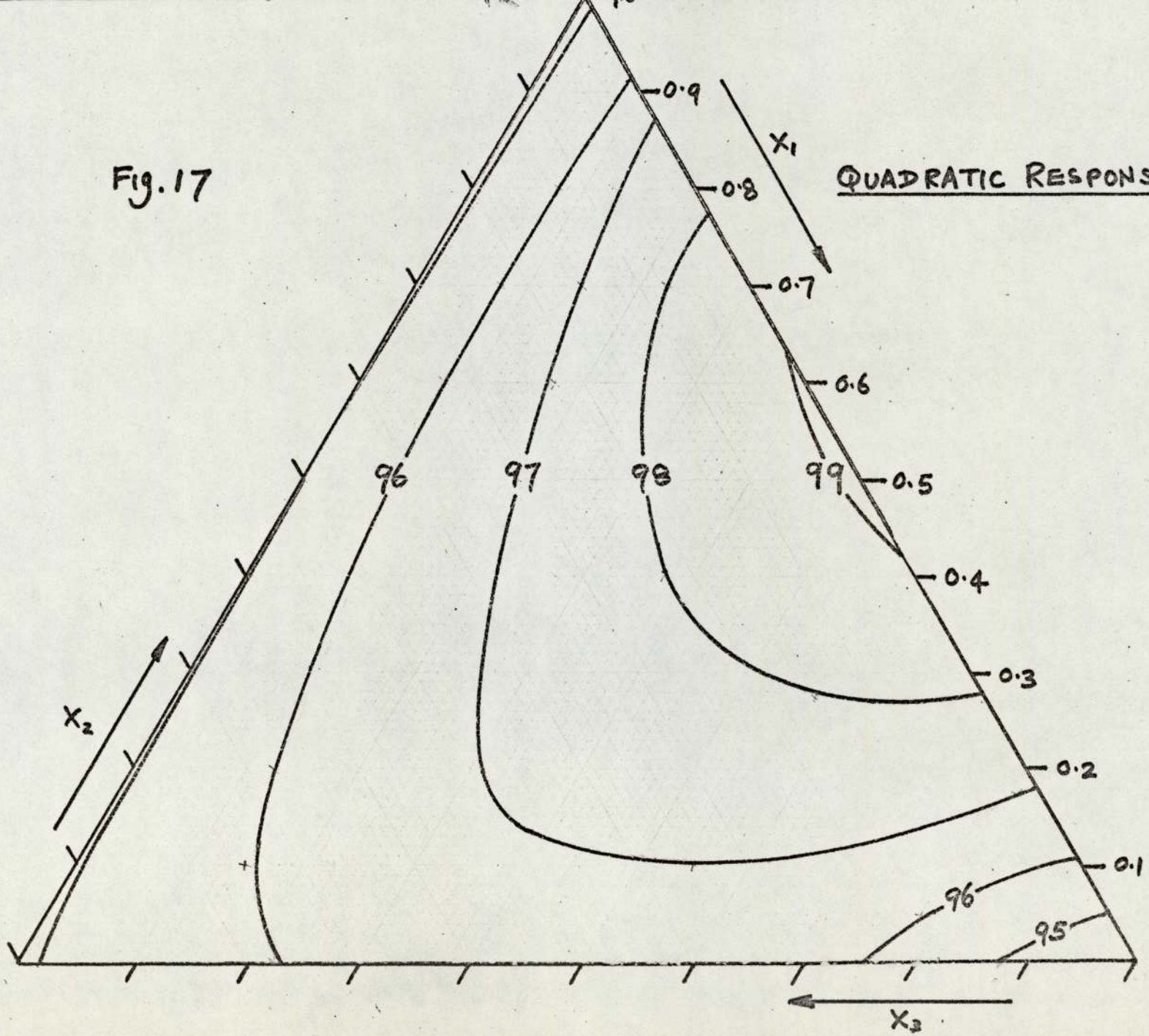
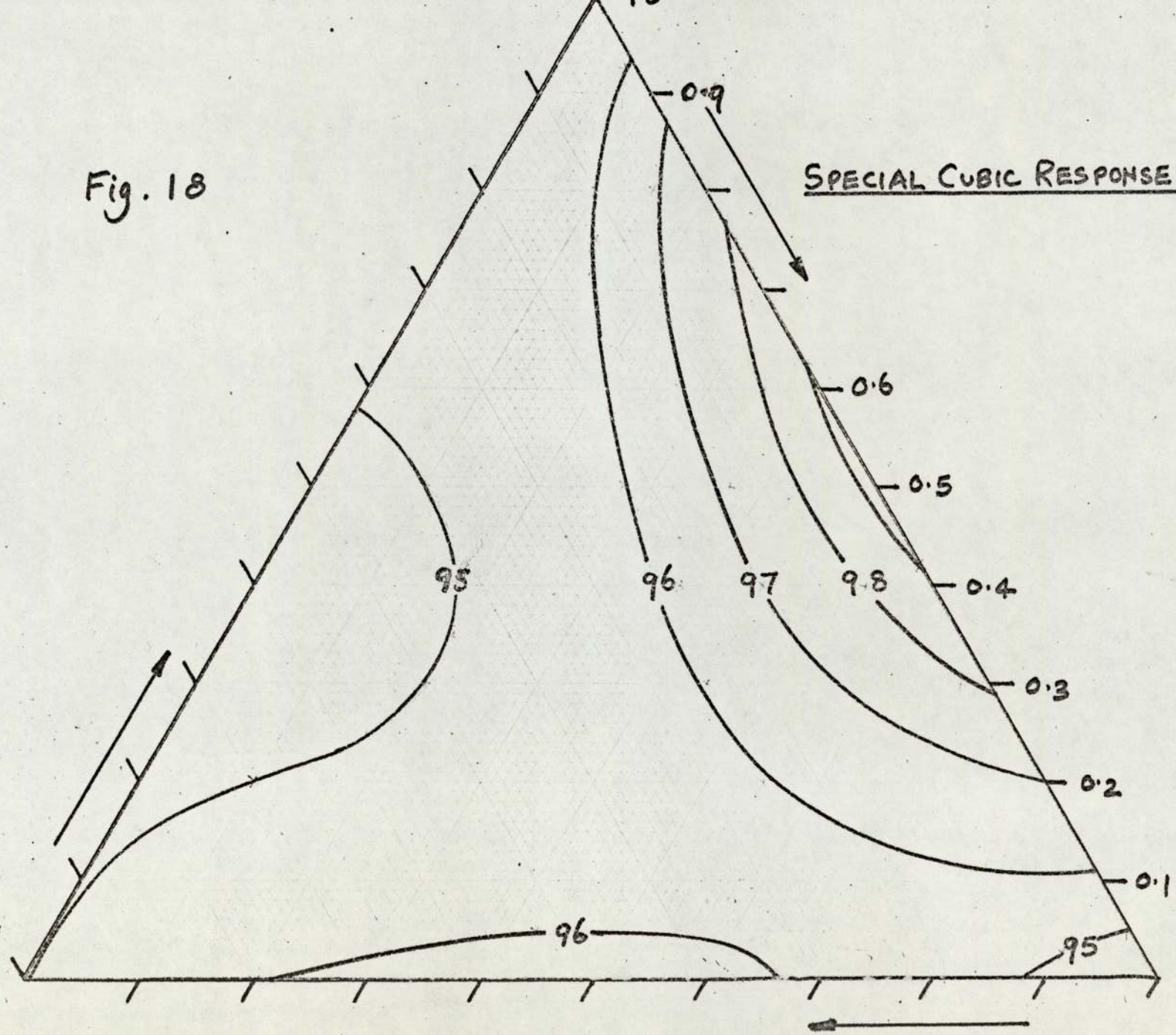


Fig. 18

SPECIAL CUBIC RESPONSE SURFACE.



When carrying out investigations which entail the examination of a response surface one is usually looking for a maximum or a minimum. If a maximum had been aimed for one would not have been led far astray and a binary mixture of equal parts of x_1 and x_3 would have been chosen. However, had one been looking for a minimum the quartic response surface would have indicated a position at the top of the valley (approx $x_1 = 0.2$, $x_2 = 0.75$, $x_3 = 0.05$) while all the others would have indicated the single component x_3. Had one been hoping to obtain some insight into the underlying mechanism one would have been led stray by using a cubic or lower degree polynomial.

It is interesting to note in passing that the writer was unable to confirm the shape of Gorman and Heinman's 94 contour. At about $x_1 = 0.3$, $x_2 = 0.6$, $x_3 = 0.1$ the sides of their 94 contour meet to form two valleys.

Evolutionary Operation (EVOP) applied to the Mixture Problem.

The technique of Evolutionary Operation developed by Box and others is now well enough known and established to require no detailed introduction. It is a management tool in which a continuous routine becomes the basic mode of operation for the plant and replaces the normal static operation. As it is intended to be used on the "shop floor" by personnel whose technical qualifications do not fit them for carrying out sophisticated statistical calculations, as many of the calculations as possible are simplified by the introduction of approximations such as the range as a measure of the standard deviation.

Consider a simple example in which, say, temperature and concentration are the two variables being investigated. Main effects and interaction are calculated in the normal way

and the range used to calculate the standard error. The significance of the effects is then assessed and a decision made concerning the way in which the variables should be altered to increase the yield.

When applying the basic ideas of EVOP to the mixture problem it is found that modifications are necessary because of the constraint $\Sigma x_i = 1$.

Suppose one had a process which, under normal operating conditions was using a mixture of three raw materials in the percentages 53%, 13%, 33% and producing a product with a property whose measured value was 106. If it is decided to carry out an EVOP investigation by changing the percentages of the mixture in the way shown in Fig. 18 , how can the figures be interpreted?

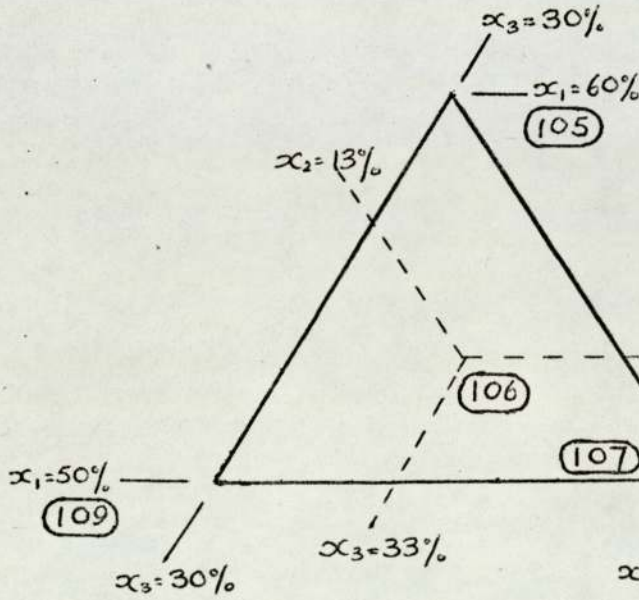


Fig. 18 Responses (ringed) for various mixture concentrations.

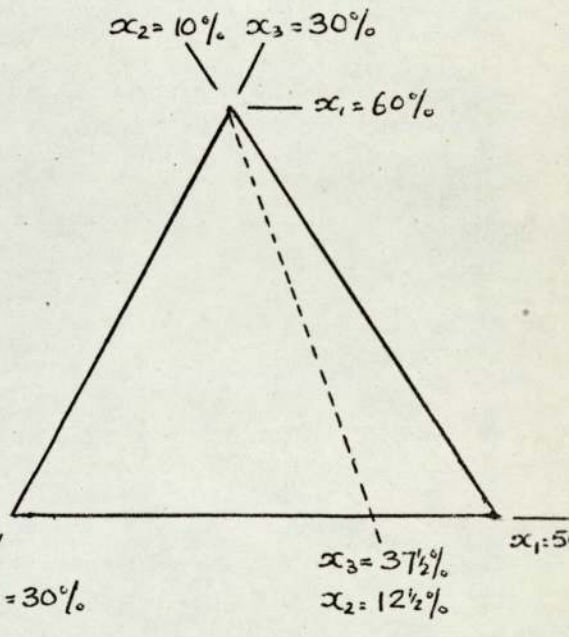


Fig. 19 $x_2 : x_3 :: 1 : 3$ along dotted line

With some experience of EVOP an investigator would initially ask what the effect of increasing or decreasing the amount of x_1 might

be. But because of the restriction $\Sigma x = 1$ he would quickly realise that any change in x_1 brings about a concomitant change in x_2 and x_3. A channel to which the writer gave some thought was to consider a line through the apex of the simplex along which the ratio $x_2:x_3$ was constant such as that shown in Fig. 19 and to investigate the effect of changing x_1 while the ratio $x_2:x_3$ was held constant and similarly for the other two pairs of ratios. However, this seemed a very contrived way of accomplishing the object and one which did not seem to be answering the question about x_1, x_2 and x_3; only about x_1 and $x_2:x_3$ etc. This approach was therefore abandoned.

Consider the simplest regression equation that could be obtained by measuring a response at the three apices of a simplex

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

The β coefficients measure the effect of the x's and it would appear sufficient to estimate the β's in order to determine the way to move up or down the response surface. This ignores the non-linear effects, but when it is realised that in EVOP one makes only small changes from the norm it can be appreciated that within the small simplex being investigated the non-linear effects will be negligible. The problem is thus resolved into finding the β's with sufficient precision to enable a decision to be made. This is accomplished by replication enabling the standard error of the coefficients to be calculated. It has been found that the inclusion in an EVOP program of the current best known conditions is important. Usually these conditions will be those at which the process is normally operating. The main advantage is the reassurance it gives to the personnel who are operating the process. Unexpected and

inexplicable variations which normally occur on the best run plants will be ascribed to the EVOP program unless this reference condition is included. Further, with this reference point a comparison can be made with the average performance achieved while running the EVOP design with the performance which would have been obtained if all the runs had been made at the reference point. This comparison is given by

(Average response over all runs of EVOP cycle)

-(average response at reference conditions)

and is called the Change in Mean Effect.

Suppose n cycles have been completed and the average response at the four points, y_o being the reference point, are as shown in Fig. 20.

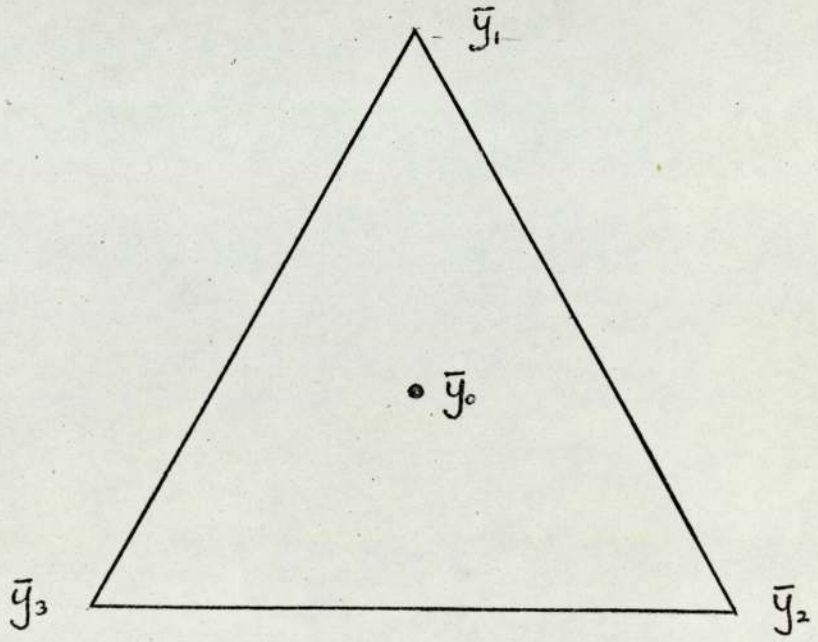


Fig. 20 :Average responses after n cycles

The change in mean is given by

$$\tfrac{1}{4}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + y_o) - y$$

$$= (\bar{y}_1 + \bar{y}_2 + \bar{y}_3 - 3y_o)/4$$

and the variance of the change in mean is given by

$$\mathrm{Var}(\tfrac{1}{4} + \tfrac{1}{4} + \tfrac{1}{4} - \tfrac{3}{4}) = 3\sigma^2/4n$$

Thus,

$$\text{S.E (change in mean)} = 0.87\sigma/\sqrt{n}$$

The standard error of the β coefficients is $\sigma/\sqrt{n}$ (See table 4 on page 48).

A further quantity which will be found useful by plant personnel is the phase mean estimated by the average of all the sets of conditions,

Phase mean = $\frac{1}{4}(\bar{y}_1 + \bar{y}_2 + \bar{y}_3 + \bar{y}_4)$

This measures the mean response over all the conditions being run in the current phase.

Student's t-distribution is not used for setting up confidence limits to test whether an effect is significant or not; for the sake of simplicity aimed at in EVOP, 2 S.E. are taken as the 95% confidence limits.

As each cycle of an EVOP program is completed the data is accumulated with data from the previous cycles to give cumulative figures for means of the responses of the various operating conditions and the standard deviation, as estimated by the range, from which the 2 S.E. limits may be calculated. The figures are entered in specially constructed tables which enable subsequent calculations to be easily as systematically calculated. A decision is then made whether to carry out another cycle using the same operating conditions or whether to move to a new phase with a new set of conditions. The figures for the mean are self-explanatory, but those involving the range need elaboration. After n cycles the following data will be available for any given operating condition,

Average after n-1 cycles	η_{n-1}
Observation on the n th.cycle	η_n
Difference, $y_n - y_{n-1}$	d_n

As the difference is the difference between one observation and the mean of n-1 previous observations all with

variance σ^2, the variance of the difference is given by

$$\text{Var}(d_n) = \sigma^2 + \frac{\sigma^2}{n-1} = \frac{n\sigma^2}{n-1}$$

so that $$\sigma = \text{Std.dev}(d_n) \left(\frac{n-1}{n}\right)^{\frac{1}{2}}$$

As mentioned above, we measure $\text{Std.dev}(d_n)$ by the range R_n and a factor W_k which is available in published tables. An estimate S_n of $\text{Std.dev}(d_n)$ is therefore provided by $R_n W_k$ leading to

$$S_n = R_n \left(\frac{n-1}{n}\right)^{\frac{1}{2}} W_k$$

$$= R_n f_{k,n}$$

where

$$f_{k,n} = \left(\frac{n-1}{n}\right)^{\frac{1}{2}} W_k$$

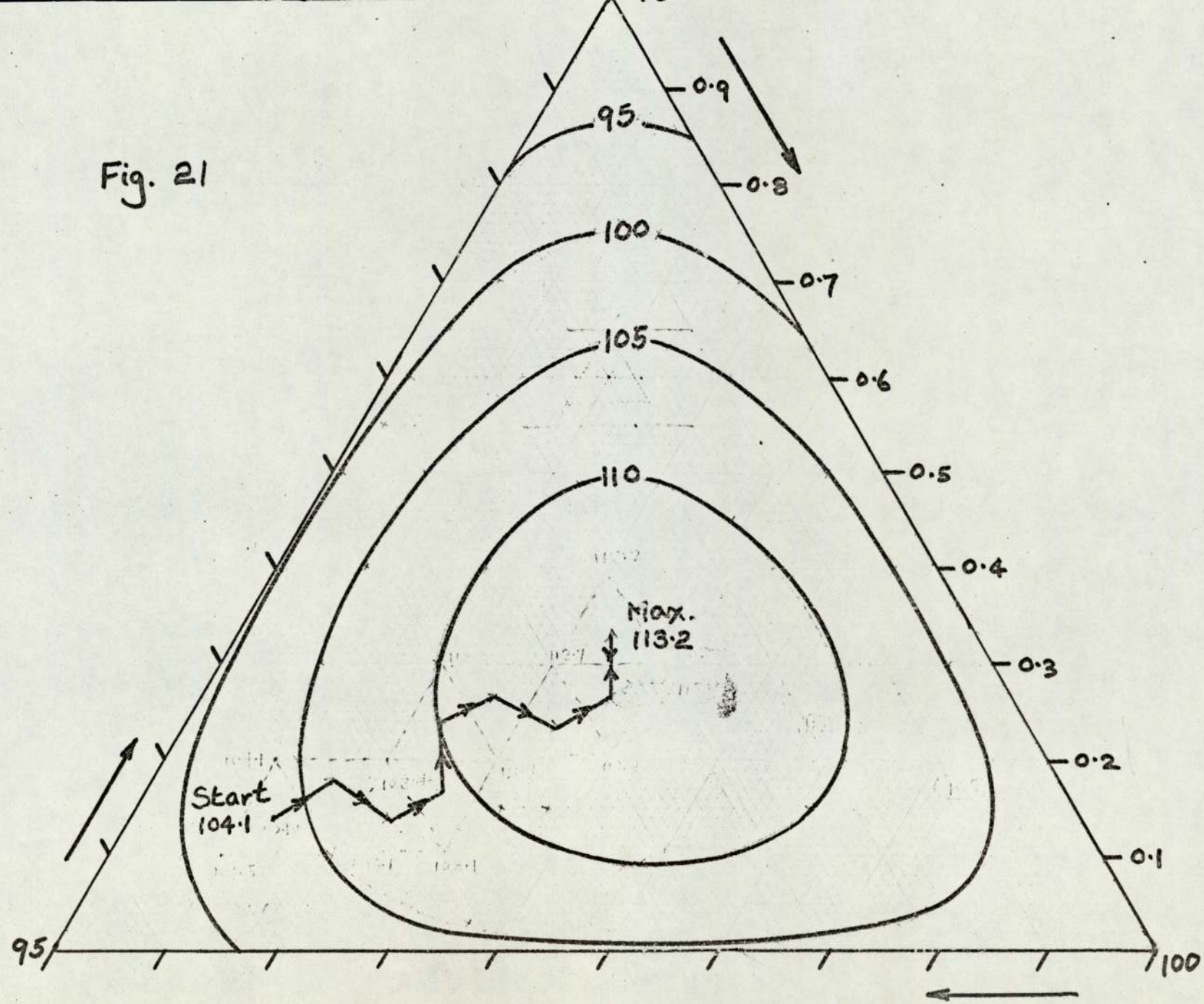
This factor may be worked out in advance and included in the EVOP calculation sheet. For example, on the fifth cycle of a program involving observations on the responses at the three corners of a simplex plus the reference point, $n = 5$ and $k = 4$. Hence, $W_4 = 0.4857$ and $f_{4,5} = \sqrt{5/(5-1)} \times 0.4299 = 0.4344$.

Having derived all the necessary equations etc. a simulation was carried out using the equation

$$y = 90x_1 + 95x_2 + 100x_3 + 27x_1x_2 + 27x_1x_2 + 27x_2x_3$$

which contains three equal terms of quadratic binary synergism and generates a surface with a single peak. The solutions to this equation were computed at 1% intervals from which Fig. 21 was drawn. It was assumed that the normal operating conditions $x_1 = 13\%$, $x_2 = 14\%$ and $x_3 = 73\%$ with a response of 104.1 units and a standard deviation of 1.0 units. To try to increase the value of the response, it was assumed that investigations were

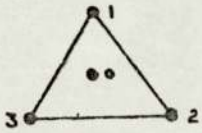
Fig. 21


95
100
105
110
0·9
0·8
0·7
0·6
0·5
0·4
0·3
0·2
0·1
Max.
113·2
Start
104·1
95
100

carried out at the points of a simplex of which the above values are the centroid or reference point (See Fig. 21). The values of the response at these points were read off from the computer print-out and a normal random deviate added to the value obtained. These figures were entered on the EVOP sheet for cycle 1, phase 1. The standard deviation (assumed to be unknown) cannot be estimated from the data of a single cycle so the entries under "Calculation of Standard Deviation" are left blank except that the prior estimate of the standard deviation (assumed to be available from previous plant operating records) is inserted in the first line. As soon as the figures from cycle 2 are available we can form an estimate of σ based on the range calculated from the two asterisked extreme values.

The 2 S.E. limits decrease as one would anticipate until cycle 4 when an "unfortunately" high random deviate increased the standard error to 1.02. This again happened in cycle 5 with an increase to 1.81. Cycle 6 compensated and the figure was reduced to 0.85. The same pattern applied in subsequent cycles and anticipated reductions in the standard error did not materialise. For didactic reasons ten simulated cycles were run before it was felt that a practical appraisal of the results should be made.

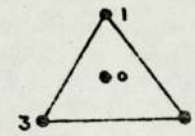
At the tenth cycle $\bar{y}_1$ and $\bar{y}_2$ are not significantly different as judged by the 2 S.E. test. However, $\bar{y}_3$ is significantly different from either $\bar{y}_1$ or $\bar{y}_2$ and at cycle 2 is the smallest of the three. It is reasonable to assume that one should move in a direction opposite to $\bar{y}_3$ for the second phase which would be centred on $x_1 = 17\%$, $x_2 = 16\%$, $x_3 = 67\%$. Proceeding in this fashion one would follow the zig-zag path shown in Fig. 21 until in the vicinity of the maximum at $x_1 = 30\%$, $x_2 = 35\%$,



CYCLE n = 1 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ 1·2
1 PREVIOUS CYCLE SUM					PREVIOUS SUM S =
2 PREVIOUS CYCLE AVERAGE					PREVIOUS AVERAGE S =
3 NEW OBSERVATIONS	104·6	103·8	105·0	101·5	NEW S = RANGE × $f_{4,n}$ =
4 DIFFERENCE 2-3					RANGE =
5 NEW SUMS 1+3	104·6	103·8	105·0	101·5	NEW SUM S =
6 NEW AVERAGES: y_i	104·6	103·8	105·0	101·5	NEW AVERAGE S = $\frac{\text{NEW SUM S}}{n-1}$ =

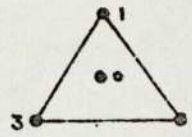
CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 103·7	
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 103·8	FOR NEW AVERAGES = $\pm\frac{2s}{\sqrt{n}}$ = ± 2·4
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 105·0	FOR NEW EFFECTS = $\pm\frac{2s}{\sqrt{n}}$ = ± 2·4
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·5	FOR CHANGE IN MEAN = $\pm\frac{1.74s}{\sqrt{n}}$ = ± 2·09
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −0·9	



CYCLE n = 2 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ 1·2
1 PREVIOUS CYCLE SUM	104·6	103·8	105·0	101·5	PREVIOUS SUM S =
2 PREVIOUS CYCLE AVERAGE	104·6	103·8	105·0	101·5	PREVIOUS AVERAGE S =
3 NEW OBSERVATIONS	105·2	105·2	104·5	101·1	NEW S = RANGE × $f_{4,n}$ = 0·73
4 DIFFERENCE 2-3	−1·6*	−1·4	+0·5*	+0·4	RANGE = 2·1
5 NEW SUMS 1+3	209·8	209·0	209·5	202·6	NEW SUM S = 0·73
6 NEW AVERAGES: y_i	104·9	104·5	104·7	101·3	NEW AVERAGE S = $\frac{\text{NEW SUM S}}{n-1}$ = 0·73

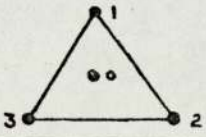
CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 104·0	
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 104·5	FOR NEW AVERAGES = $\pm\frac{2s}{\sqrt{n}}$ = ± 1·7
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 104·7	FOR NEW EFFECTS = $\pm\frac{2s}{\sqrt{n}}$ = ± 1·7
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·3	FOR CHANGE IN MEAN = $\pm\frac{1.74s}{\sqrt{n}}$ = ± 1·5
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −1·2	



CYCLE n = 4 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.	
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ	1·2
1 PREVIOUS CYCLE SUM	314·7	313·0	315·2	303·6	PREVIOUS SUM s	= 1·25
2 PREVIOUS CYCLE AVERAGE	104·9	104·3	105·1	101·2	PREVIOUS AVERAGE s	= 0·63
3 NEW OBSERVATIONS	104·3	102·2	105·1	101·5	NEW s = RANGE $\times f_{4,n}$	= 1·01
4 DIFFERENCE 2–3	0·6	2·1*	0·0	−0·3*	RANGE	= 2·4
5 NEW SUMS 1+3	419·0	415·2	420·3	405·1	NEW SUM s	= 3·06
6 NEW AVERAGES: y_i	104·8	103·8	105·1	101·3	NEW AVERAGE s = $\frac{\text{NEW SUM } s}{n-1}$	= 1·02

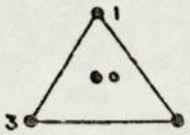
CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 103·3	
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 103·8	FOR NEW AVERAGES = $\pm\frac{2s}{\sqrt{n}}$ = ± 1·02
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 105·1	FOR NEW EFFECTS = $\pm\frac{2s}{\sqrt{n}}$ = ± 1·02
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·3	FOR CHANGE IN MEAN = $\pm\frac{1{\cdot}74s}{\sqrt{n}}$ = ± 0·89
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −1·0	



CYCLE n = 6 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.	
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ	1·2
1 PREVIOUS CYCLE SUM	523·3	518·7	523·4	505·4	PREVIOUS SUM s	= 4·05
2 PREVIOUS CYCLE AVERAGE	104·7	103·7	104·7	101·1	PREVIOUS AVERAGE s	= 1·81
3 NEW OBSERVATIONS	105·1	103·9	104·6	101·0	NEW s = RANGE $\times f_{4,n}$	= 0·22
4 DIFFERENCE 2–3	−0·4*	−0·2	0·1*	0·1	RANGE	= 0·5
5 NEW SUMS 1+3	628·4	622·6	627·0	606·5	NEW SUM s	= 4·27
6 NEW AVERAGES: y_i	104·7	103·8	104·5	101·1	NEW AVERAGE s = $\frac{\text{NEW SUM } s}{n-1}$	= 0·85

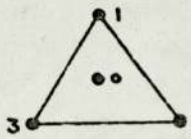
CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 103·6	
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 103·8	FOR NEW AVERAGES = $\pm\frac{2s}{\sqrt{n}}$ = ± 0·70
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 104·5	FOR NEW EFFECTS = $\pm\frac{2s}{\sqrt{n}}$ = ± 0·70
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·1	FOR CHANGE IN MEAN = $\pm\frac{1{\cdot}74s}{\sqrt{n}}$ = ± 0·60
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −1·5	



CYCLE n = 9 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.	
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ	1·2
1 PREVIOUS CYCLE SUM	837·8	830·0	836·0	808·7	PREVIOUS SUM S	= 6·25
2 PREVIOUS CYCLE AVERAGE	104·7	103·7	104·5	101·1	PREVIOUS AVERAGE S	= 0·88
3 NEW OBSERVATIONS	104·0	104·9	104·9	101·1	NEW S = RANGE $\times f_{4,n}$	= 0·87
4 DIFFERENCE 2-3	0·7*	−1·2*	−0·4	0·0	RANGE	= 1·9
5 NEW SUMS 1+3	941·8	934·9	940·9	909·8	NEW SUM S	= 7·12
6 NEW AVERAGES: $\bar{y}_i$	104·7	103·7	104·6	101·1	NEW AVERAGE S = $\frac{\text{NEW SUM S}}{n-1}$	= 0·85

CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.	
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 103·7		
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 103·7	FOR NEW AVERAGES = $\frac{\pm 2s}{\sqrt{n}}$	= ± 0·57
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 104·6	FOR NEW EFFECTS = $\frac{\pm 2s}{\sqrt{n}}$	= ± 0·57
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·1		
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −1·0	FOR CHANGE IN MEAN = $\frac{\pm 1.74s}{\sqrt{n}}$	= ± 0·52



CYCLE n = 10 PHASE = 1

CALCULATION OF AVERAGES					CALCULATION OF STD. DEV.	
OPERATING CONDITIONS	(0)	(1)	(2)	(3)	PRIOR ESTIMATE OF σ	1·2
1 PREVIOUS CYCLE SUM	941·8	934·9	940·9	909·8	PREVIOUS SUM S	= 7·12
2 PREVIOUS CYCLE AVERAGE	104·7	103·7	104·6	101·1	PREVIOUS AVERAGE S	= 0·85
3 NEW OBSERVATIONS	103·6	103·9	103·9	99·7	NEW S = RANGE $\times f_{4,n}$	= 0·74
4 DIFFERENCE 2-3	1·1	−0·2*	0·7	1·4*	RANGE	= 1·6
5 NEW SUMS 1+3	1045·4	1038·8	1044·8	1009·5	NEW SUM S	= 7·86
6 NEW AVERAGES: $\bar{y}_i$	104·5	103·9	104·5	100·9	NEW AVERAGE S = $\frac{\text{NEW SUM S}}{n-1}$	= 0·91

CALCULATION OF EFFECTS		CALCULATION OF S.E. LIMITS.	
PHASE MEAN = $\frac{1}{4}(\bar{y}_0 + \bar{y}_1 + \bar{y}_2 + \bar{y}_3)$	= 103·5		
β_1 (COEFFICIENT x_1) = $\bar{y}_1$	= 103·9	FOR NEW AVERAGES = $\frac{\pm 2s}{\sqrt{n}}$	= ± 0·52
β_2 (COEFFICIENT x_2) = $\bar{y}_2$	= 104·5	FOR NEW EFFECTS = $\frac{\pm 2s}{\sqrt{n}}$	= ± 0·52
β_3 (COEFFICIENT x_3) = $\bar{y}_3$	= 101·0		
CHANGE IN MEAN EFFECT = PHASE MEAN − $\bar{y}_0$	= −1·0	FOR CHANGE IN MEAN = $\frac{\pm 1.74s}{\sqrt{n}}$	= ± 0·53

x_3 = 35% with η = 113.2. However, before this point was reached one would find that the amount of effort involved in running the scheme for the small increase obtained was not justified from an economic point of view and the program would be terminated before the true optimum was reached.

If the response surface is examined it is obvious why $\bar{y}_1$ and $\bar{y}_2$ do not become significantly different at even the tenth cycle; they both lie approximately on the same contour. Harrington has found this same effect with conventional square and cube EVOP patterns when applied to chemical processes and attributes this to the too cautious choice of the limits between which the variables were changed. He suggests that if the first few cycles do not reveal any change a larger pattern rotated at 45° to the first should be selected as shown in Fig. 22.

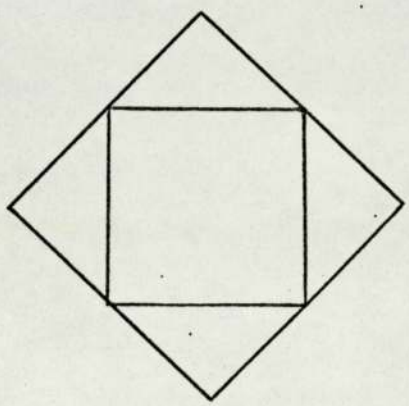


Fig. 22

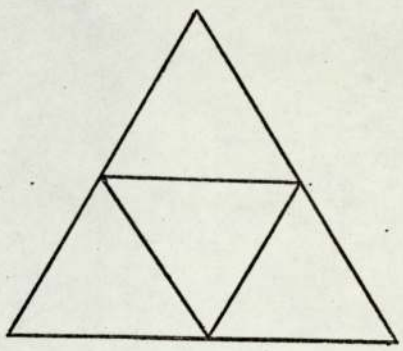


Fig. 23

The effect of this is to double the area of the space being examined. If one extrapolates the technique to a simplex the area is quadrupled and this might be too bold a step resulting in the production of substandard material. The writer

suggests two alternatives; rotate the simplex through 60°. This will invert it and if by chance two of the points do lie on the same contour line the defect will be corrected (Fig. 24). Alternatively one can double the area of the simplex by changing its height from $\sqrt{3}$ units to $\sqrt{6}$. This amounts to increasing the height of the simplex by 41.6% or, for practical purposes, two fifths.

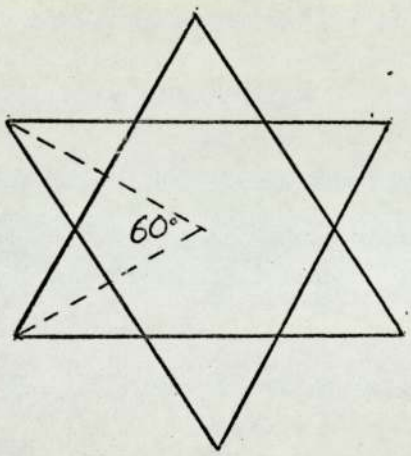


Fig. 24

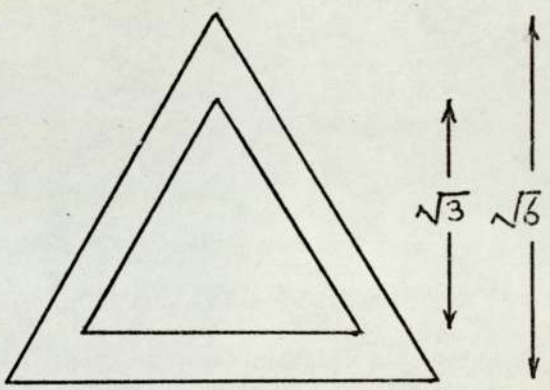


Fig. 25

Practical Experience with the Technique.

These techniques have been used by the writer to investigate several three component systems. A (3,3) simplex lattice was used in all cases. One investigation was to examine the effect of two different types of polyacrylamide resins and tetrasodium pyrophosphate (T.S.P.P) as a bonding agent for foundry sands.

Figs. 26 to 28 give the response surfaces for different physical properties of sand bonded with different resin/T.S.P.P. mixtures. As high values are required for all

Fig. 26

SHATTER INDEX OF SAND BONDED WITH A MIXTURE OF POLYMERS.

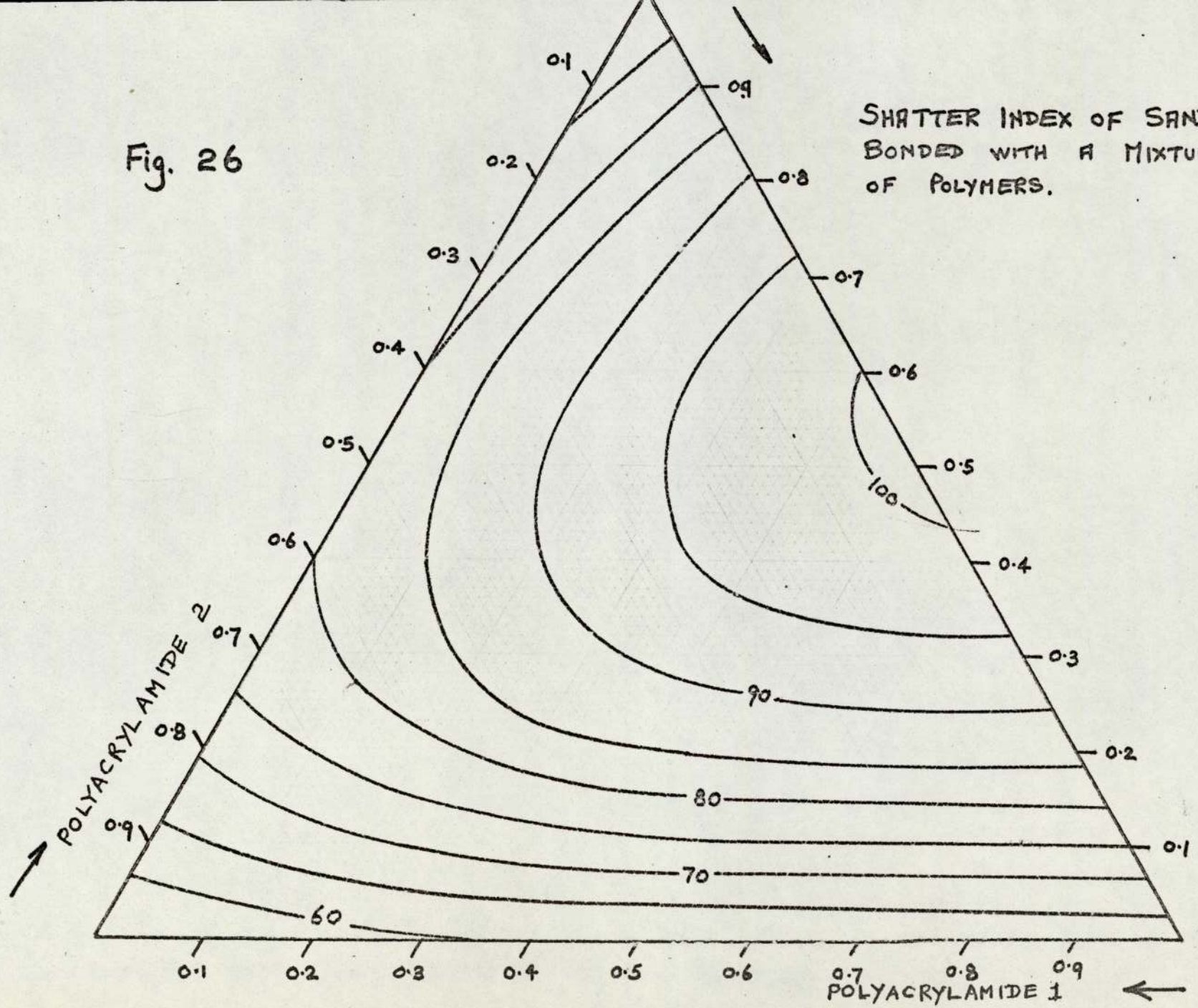


Fig 27

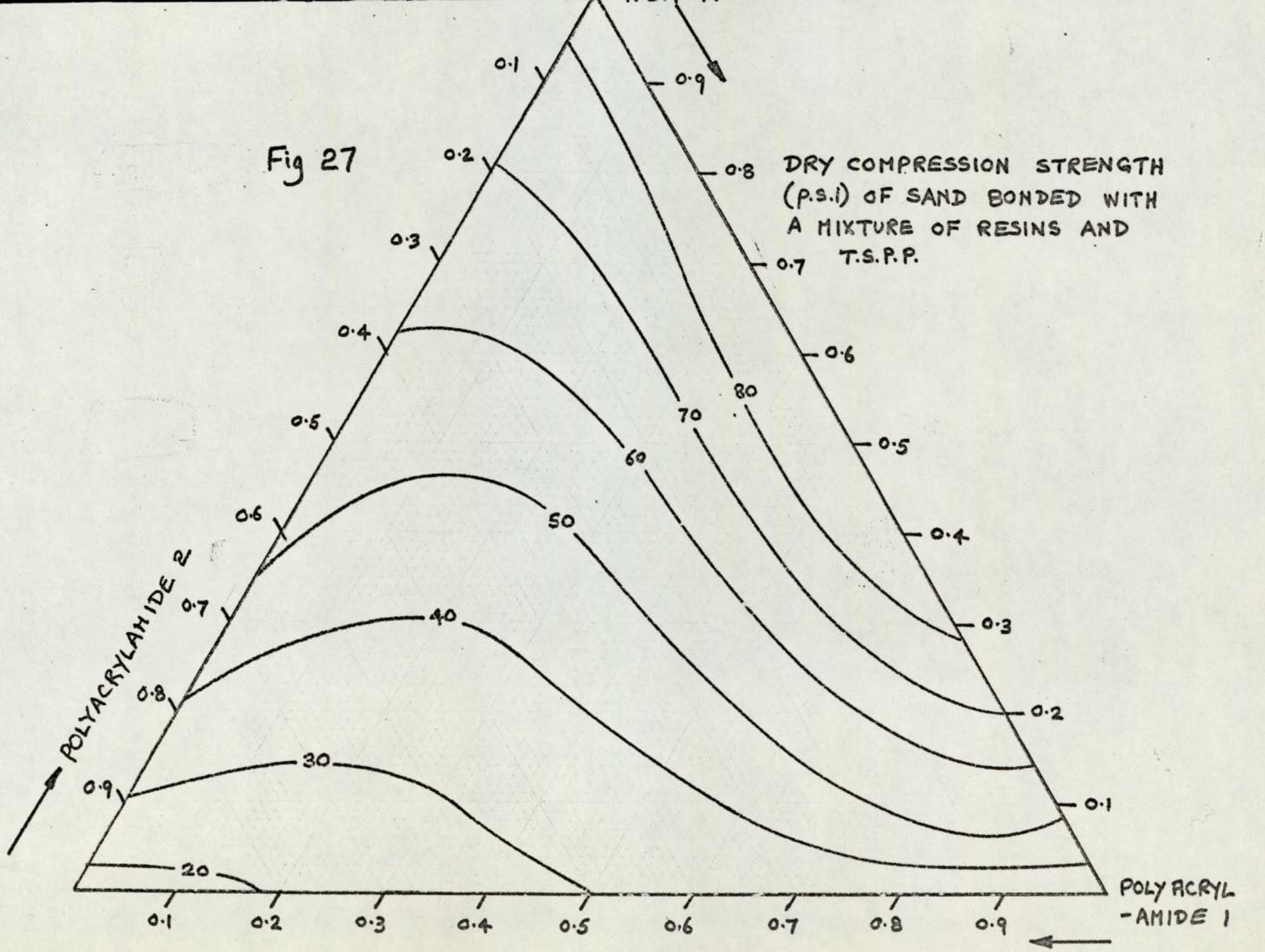

DRY COMPRESSION STRENGTH (p.s.i) OF SAND BONDED WITH A MIXTURE OF RESINS AND T.S.P.P.
POLYACRYLAMIDE 2
POLY ACRYL -AMIDE 1
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
20
30
40
50
60
70
80

Fig 28

GREEN TENSILE STRENGTH (g/cm²) OF SAND BONDED WITH MIXTURES OF POLYMERS.

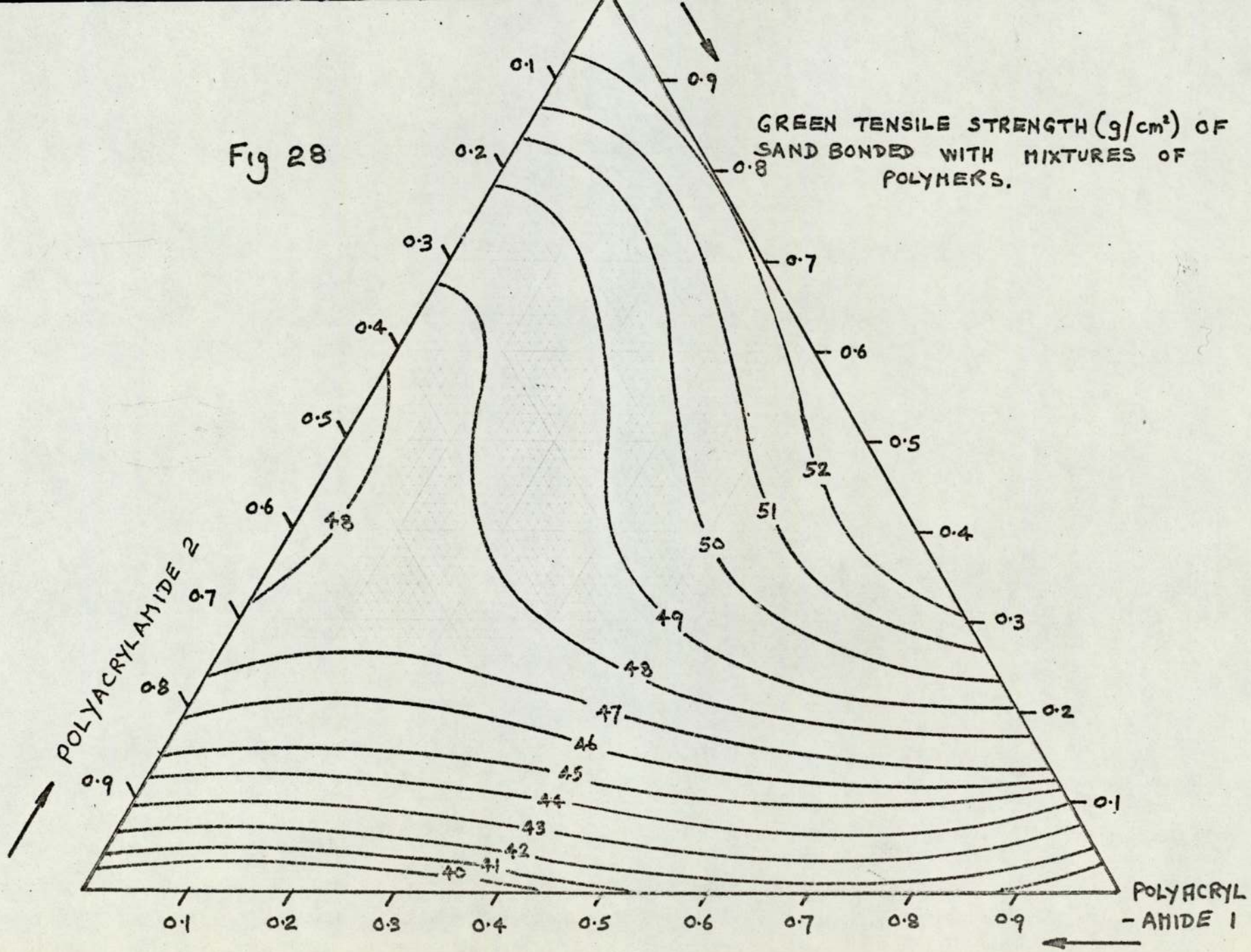
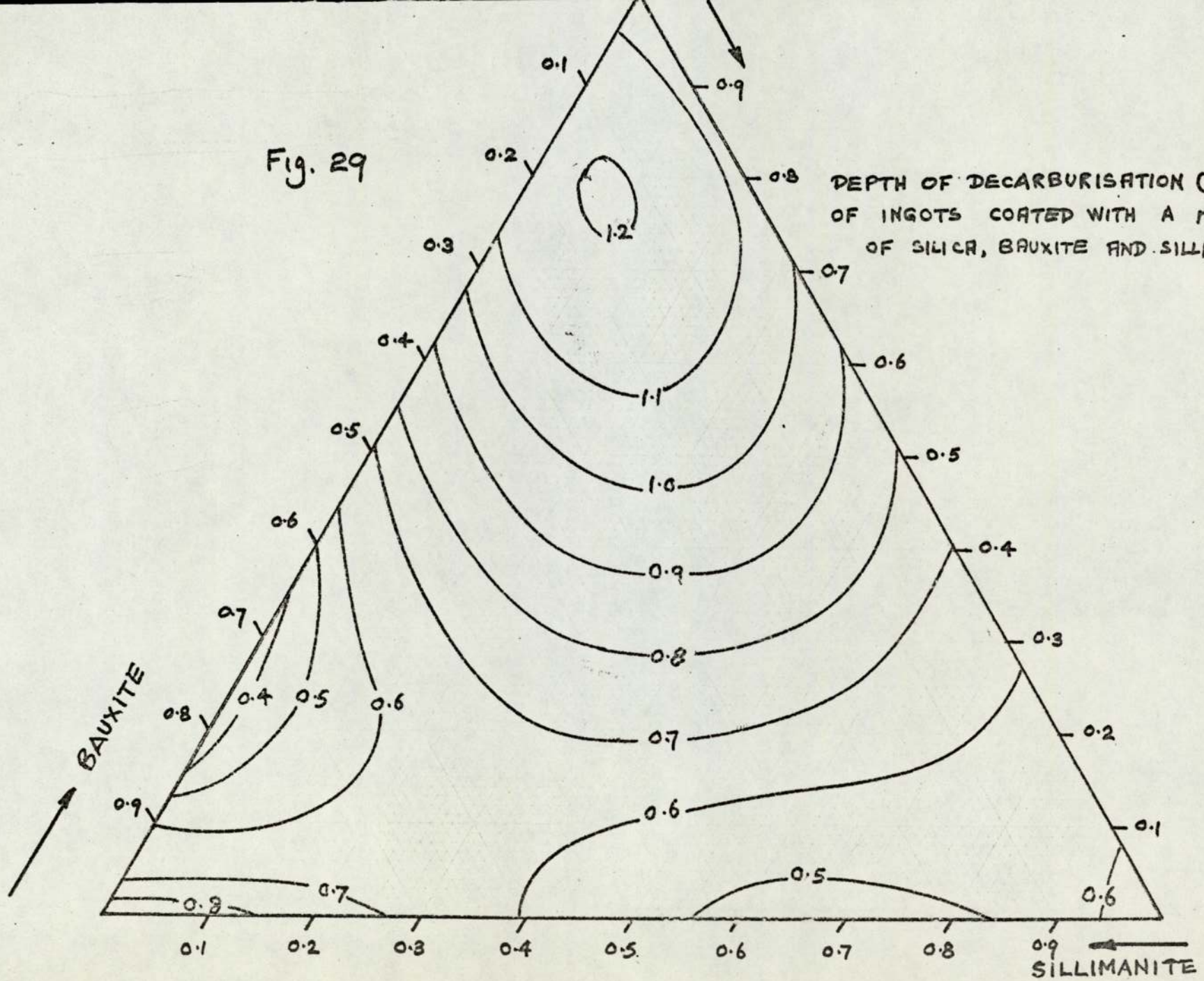


Fig. 29

DEPTH OF DECARBURISATION (MM) OF INGOTS COATED WITH A MIXTURE OF SILICA, BAUXITE AND SILLIMINITE.



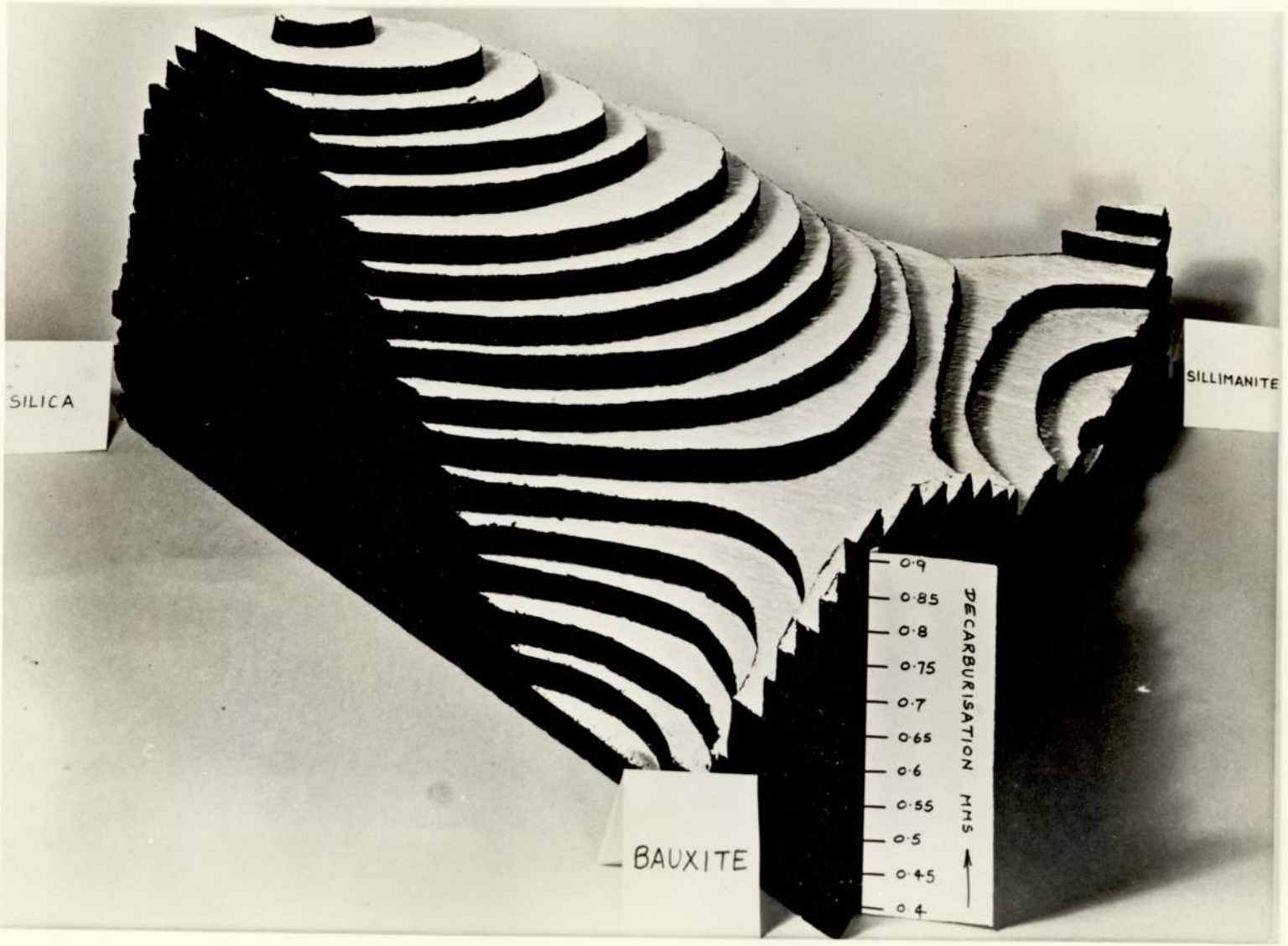
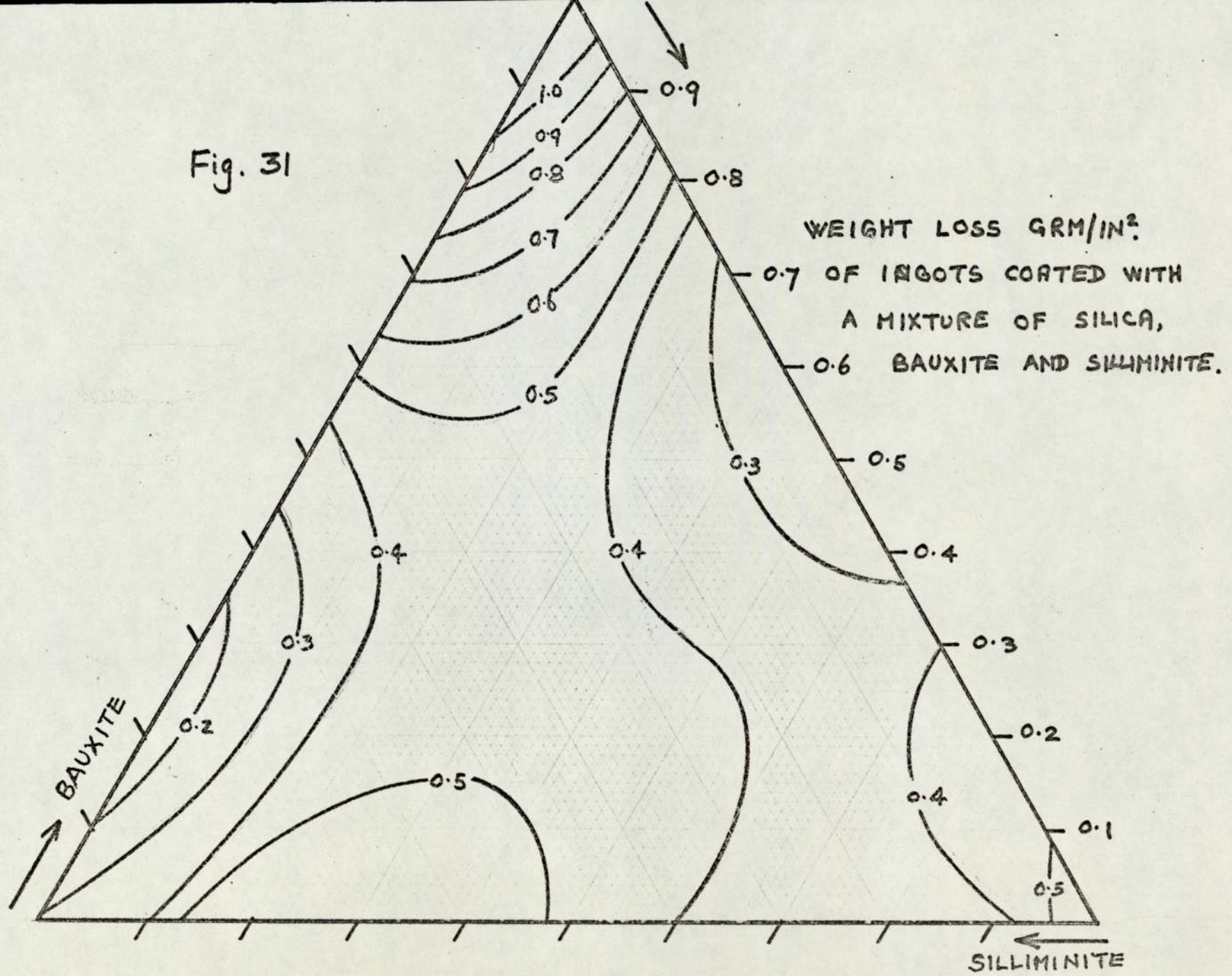


Fig. 30 Model of Response Surface of Fig. 29.

Fig. 31


WEIGHT LOSS GRM/IN².
OF INGOTS COATED WITH
A MIXTURE OF SILICA,
BAUXITE AND SILLIMINITE.
BAUXITE
SILLIMINITE
1.0
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1

four physical properties it is apparent by examining the response surfaces that polyacrylamide 2 in any combination is not acceptable and that a mixture of 70% T.S.P.P./30% polyacrylamide 1 would have the desired properties.

Another investigation concerned the development of a coating which would inhibit decarburisation of steel ingots during preheating. During this process carbon is lost from the surface layers and the properties of the metal in this region are altered. Allowance has to be made for this and wastage of metal results. From technological considerations it was felt that a paint made from silliminite, bauxite and silica would be efficacious in retarding decarburisation and a series of experiments were undertaken to test this by coating small ingots of steel and measuring the depth to which decarburisation had taken place after heating and cooling. Fig. 29 shows the response surface obtained for decarburisation and Fig. 30 is three dimensional model of the response surface. Fig. 31 is the reponse surface for the loss in weight of the ingots. The interesting point to notice is that silica or bauxite alone allow heavy decarburisation to take place (approx 1 mm.). But in combination (75% bauxite/25% silica) a coating is produced which is better than any other combination, allowing decarburisation to proceed to a depth of less than 0.4 mm. Further experiments confirmed this and as an added check a paint was made up corresponding to the position of the maximum (78% silica/10% silliminite/12% bauxite). After processing with this coating the ingot was not only heavily decarburised but pitted and scaled.

In an experiment which was carried out as part of the same investigation with aluminium/magnesia/silicon carbide mixtures a difficulty was encountered in that a stable coating could be made up only for certain combinations of the three

Fig 32

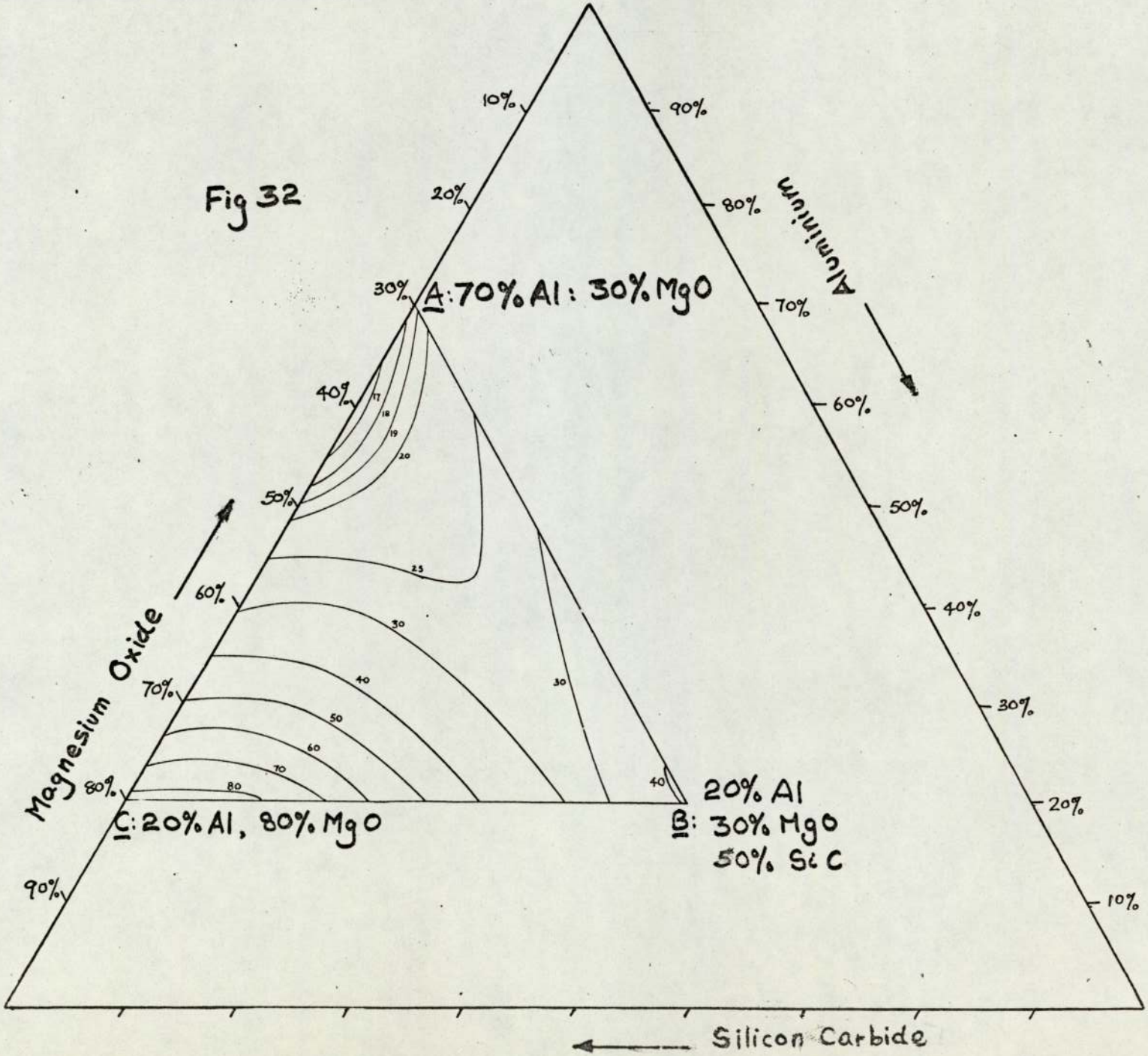

A: 70% Al : 30% MgO
B: 20% Al
30% MgO
50% SiC
C: 20% Al, 80% MgO
Magnesium Oxide
Aluminium
Silicon Carbide
10%
20%
30%
40%
50%
60%
70%
80%
90%
17
18
19
20
25
30
40
50
60
70
80

Fig 33

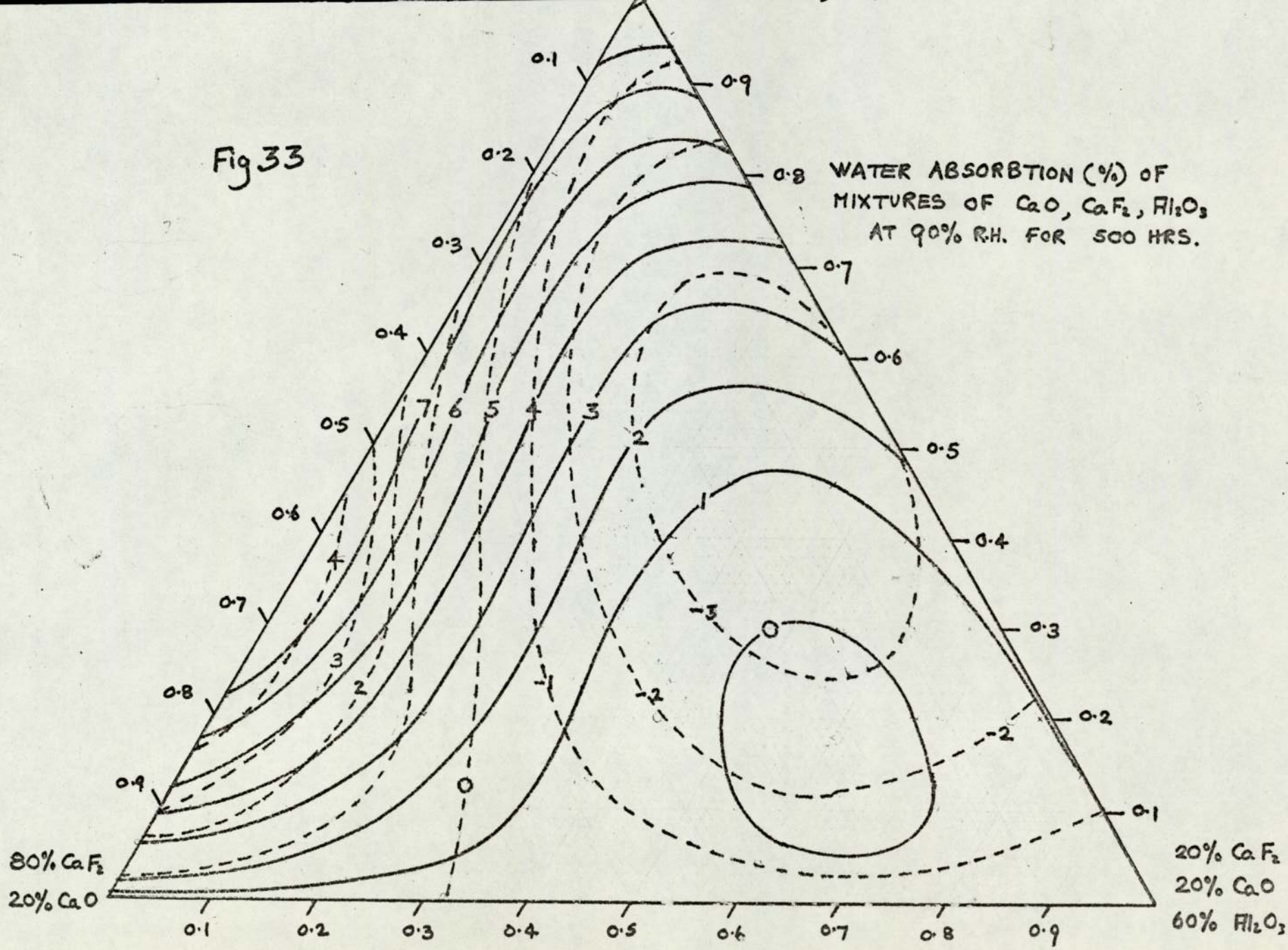
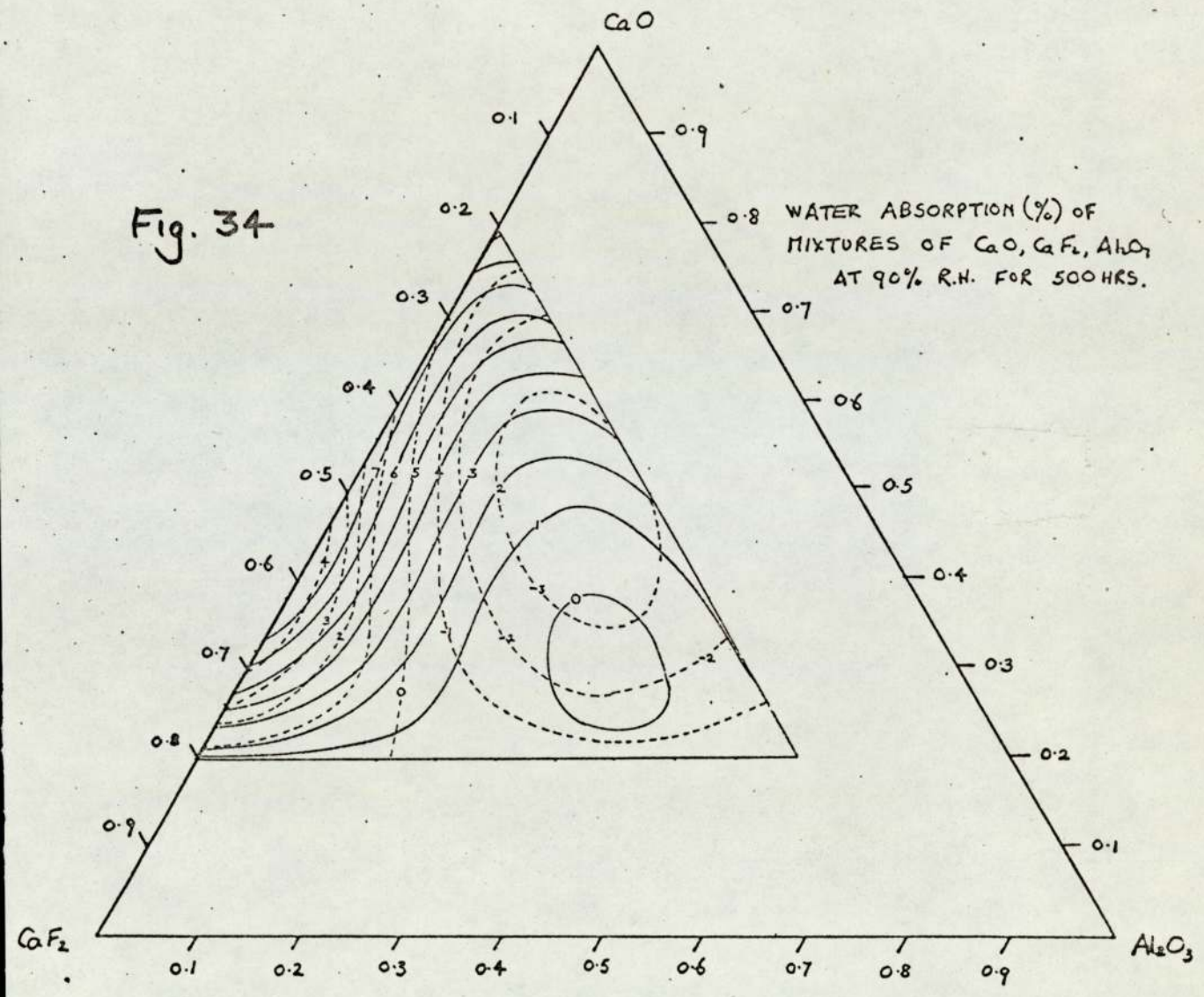

WATER ABSORBTION (%) OF MIXTURES OF CaO, CaF₂, Al₂O₃ AT 90% R.H. FOR 500 HRS.
80% CaF₂
20% CaO
20% CaF₂
20% CaO
60% Al₂O₃
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9

Fig. 34



WATER ABSORPTION (%) OF MIXTURES OF CaO, CaF_2, Al_2O_3 AT 90% R.H. FOR 500 HRS.

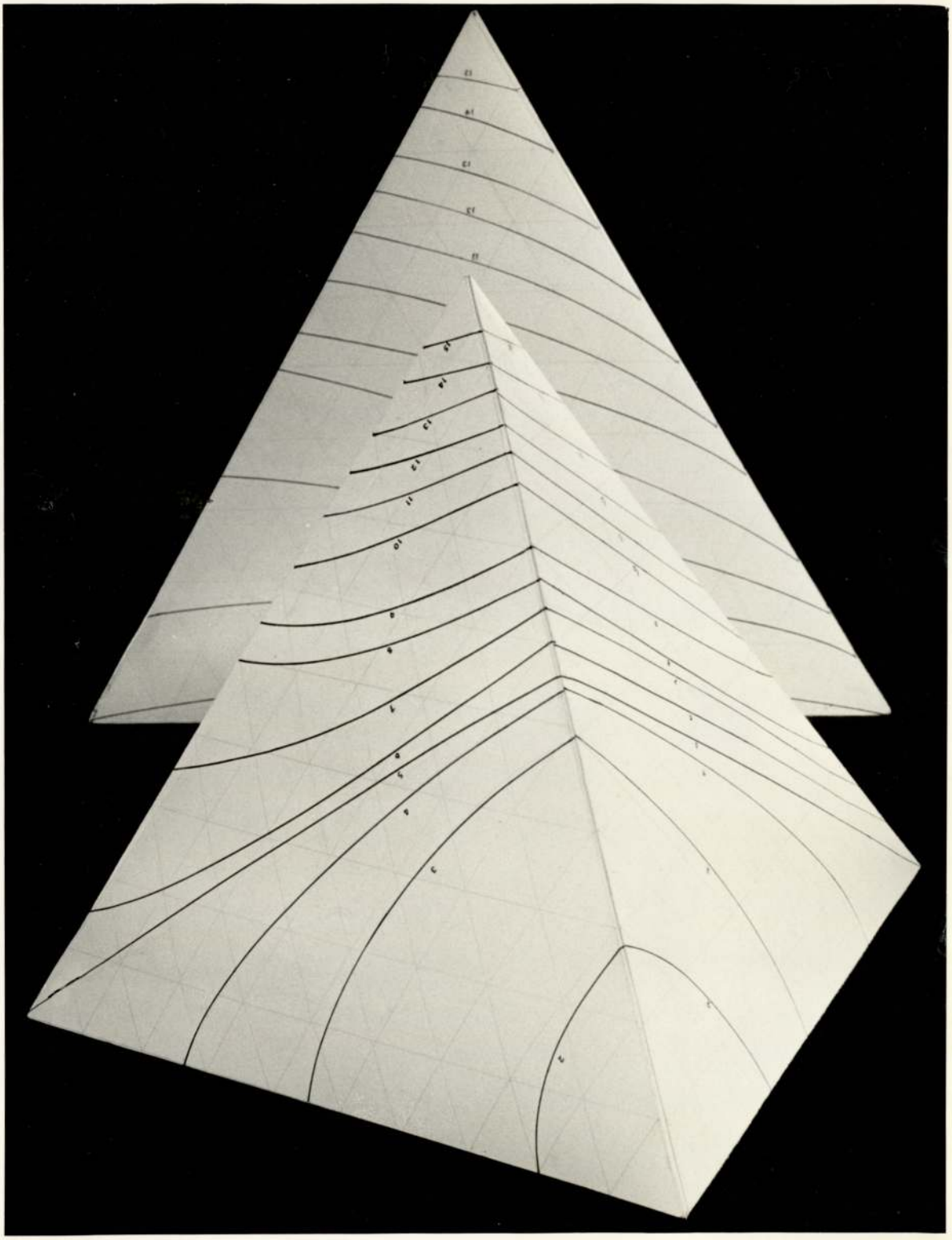


Fig.35 Response Surface for a 4 Component Mixture Showing Reflection of Third Side.

components. This meant that observations involving the pure components could not be made. This problem was solved using pseudocomponents. The mixtures A,B and C were made up (Fig.32) and these were regarded as single components for the purpose of the experiment. When the observations had been made and the response surface plotted it was reduced photographically and superimposed in the correct place on the simplex as shown.

The same technique had to be used when attempting to find the best composition for starter tablets for electro-slag refining. These tablets, which in this case were made of calcium fluoride, calcium oxide and alumina, are used to help the electric arc to strike at the beginning of the process and to form a liquid basic slag in which the metal is purified. Due to the hygroscopic nature of the calcium oxide the tablets deteriorated on storage. There was a certain latitude about the composition of the tablets and an investigation was put in hand to find the composition which was the least hygroscopic. Pseudocomponents were used again, this time because mixtures outside certain bounds were of no technological interest as it was known that they would not work. The response surface of the pseudocomponents and its synergism and antagonism are shown in Fig.33 while the response surfaces to the correct scale is shown in Fig.34. The optimum composition corresponds to about equal proportions of the three components.

Fig. 35 shows a model of a response surface for a four component system. The four components are edible oils and the response is a measure of quality.

Discussion.

Scheffe's basic experimental designs are the simplex

lattice and the simplex centroid. In the former we are completely free to choose the degree of the polynomial we wish to fit while in the latter the degree is always equal to the number of components involved except when the designs are fractionated.

Box and Draper (1959) and Box and Hunter (1957) suggest that when a function was to be graduated by a polynomial, suitable requirements for a response surface design are as follows, although to be of value for a specific purpose a design will not need to possess them all.

a) The design should allow the approximating polynomial of degree n to be estimated with satisfactory accuracy within the region of interest.

b) It should allow a check to be made on the representational accuracy of the assumed polynomial.

c) It should not contain an excessively large number of experimental points.

d) It should lend itself to "blocking"

e) It should form a nucleus from which a satisfactory design of order n+1 can be built in case the assumed degree of the polynomial proves inadequate.

In Scheffe's designs the polynomials obtained utilise all the experimental points and pass through them all. By taking a larger number of observational points a better approximation to the true function may be obtained. The only limit is the number of observations one can physically achieve. Requirement (a) is thus satisfied.

In order to test the accuracy of the polynomial Scheffe uses check points and applies a t-test to test the significance of the difference between the observed response at that point and the response predicted by the polynomial. This does not satisfy

some authorities who feel that a deficiency in Scheffe's work is that he does not consider the variance of the predictions but only the variances of the estimated coefficients. However, in view of the fact that a check can be made, albeit not to everyone's satisfaction, we can say that requirement (b) is satisfied.

It seems to the writer that what constitutes an "excessively large number of points" very much depends upon the experiments. The experiment described on page 103 regarding the development or a sand bonding agent was very easily carried out. Sand cores for testing could be produced at the rate of 3-4 per minute and a batch of 20 could be made at a time. A cubic model involving ten observation points, each point replicated 10 times was not considered excessive. On the other hand in an investigation involving the casting of 5 ton ingots followed by sawing in half, smoothing, polishing etc.any number in excess of two was considered too many!

However, there is no doubt that Lambrakis's q-component mixture designs do lead to a large number of points and suffer from the further restriction that they are undefinable when the number of components is three or less.

Simplex lattice and simplex centroid designs do not lend themselves to blocking. Box and Hunter (1957) describe how rotatable designs can be broken down into blocks. E.g. A hexagonal design can be regarded as two equilateral triangles; an octangoral design can be regarded as two squares while a nonagonal design can be regarded as three equilateral triangles etc. (Fig.36).

Scheffe's simplexes cannot be split up in this way because the points are not equiradially distributed about the centroid.

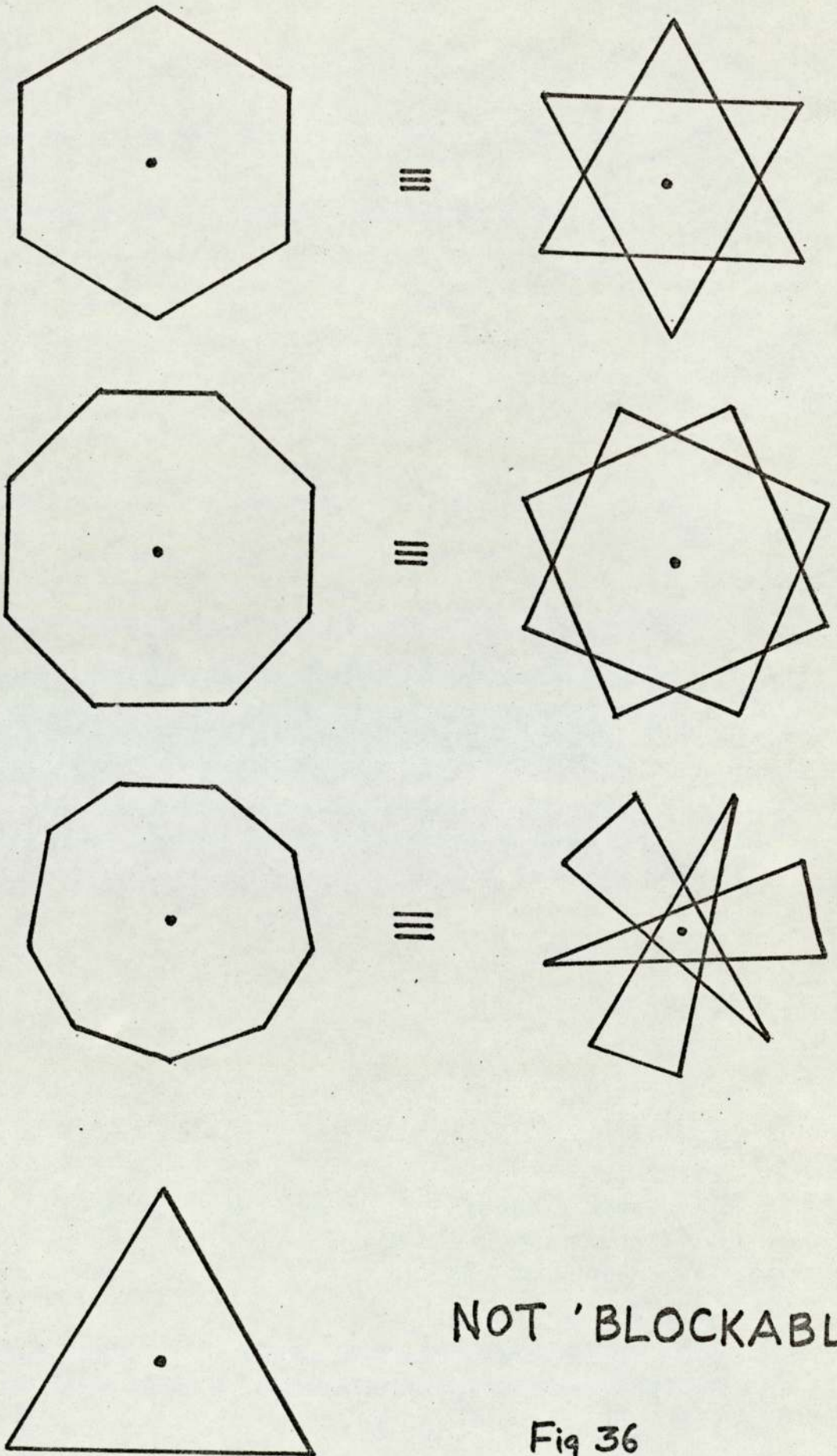


Fig 36

We have seen how a special cubic model can be built up from the quadratic model. This step, at least, fulfils requirement (e). However, the special cubic has only four points in common with the cubic model and to use the former as a stepping stone to the latter would require the rejection of three observations and the making of four more at different coordinates. If one examines Fig. 2 it can be appreciated that a (q,m) lattice is a poor stepping stone to a (q,m+1) lattice but a better one to a (q,m+2) lattice. Scheffe's simplex lattice designs do not fulfil requirement (e) completely. The centroid designs do not fulfil it at all as degree of the polynomial is determined by the number of components.

Draper and Lawrence (1965) point out another disadvantage of Scheffe's designs: the characteristics of a mixture of three components are not exhibited by mixtures which do not contain all the components. Consider for a two component mixture the petrol-oil mixture used in two-stroke engines. Experiments on "all petrol-no oil" or "no petrol-all oil" would not exhibit results typical of the mixture.

All in all, Scheffe's designs would seem to be not altogether satisfactory by the criteria given by Box et al. However, their six criteria may be the council of perfection.

Turning now to Draper and Hunter's designs; for any polynomial there are an infinite number of designs.

For the case where a first degree polynomial is to be fitted and we wish to guard against the possibility that the true function is a second degree polynomial (i.e. $d_1 = 1$ and $d_2 = 2$) nine designs have been calculated and for the case $d_1 = 2$, $d_2 = 3$, twenty two designs have been calculated.*

---

*The corresponding figures for 4 components are ten and twenty-six

Some of the designs are not accurate and only those with discrepancies less than 10% are given. If an exact design is required one must take at least 12 observations. Requirement a) is thus satisfied to a certain extent, but in opposition to requirement b).

With these designs an analysis of variance can be carried out thus fulfilling requirement b). The writer feels that this is of great weight when compared with Scheffe's contrivance of using check points.

However, the designs are not blockable any more than are those of Scheffe.

Requirement e) is not fulfilled although one can pick out one design where it nearly is. For $d_1=1$, $d_2=2$ there exists one design where $p = 0.733$ and $q = 0.457$ and for $d_1=2$, $d_2=3$ there is a design where $p_1 = 0.748$, $p_2 = 0.445$, $q_1 = 0.468$ and $q_2 = 0.156$. This seems to be the only one where one might feel that the difference is small enough to be ignored for practical purposes.

Some more specific points will now be discussed. In the first instance the relative merits of simplex lattice and simplex centroid designs.

Estimation is more difficult with the (q,m) lattice designs. Equations for the coefficients of the general polynomial are available only for $m \leq 4$ while the coefficients for the centroid designs are easily obtained from the formula given on page 35 . Further, the centroid designs are easily fractionated by deleting terms while this technique applied to lattice designs does not seem to be satisfactory. For instance if one decides to reduce the number of observations on the cubic model by deleting all terms ≥ 3; as well as deleting $\beta_{ijk} x_i x_j x_k$ one should logically delete the terms $\gamma_{ij} x_i x_j (x_i - x_j)$. Similarly for a quartic model; if the terms $\beta_{ijk\ell} x_i x_j x_k x_\ell$ goes so should the terms

$\delta_{ij}x_ix_j(x_i-x_j)^2$, $\beta_{iijk}x_i^2x_jx_k$, $\beta_{ijjk}x_ix_j^2x_k$ and $\beta_{ijkk}x_ix_jx_k^2$.

The simplex lattice permits more asymmetry in the regression function fitted. Along the edge of a simplex centroid design the regression function reduces to

$$y = \beta_i x_i + \beta_j x_j + \beta_{ij} x_i x_j.$$

But for a simplex lattice design the function is

$$y = \beta_i x_i + \beta_j x_j + \beta_{ij} x_i x_j + \gamma_{ij} x_i x_j (x_i - x_j)$$

which permits a cubic term. The centroid design can show only a synergism which is symmetric about the edge, while the lattice design can show asymmetric synergisms.

When process variables are included the technique which was illustrated for the centroid designs will work equally well for lattice designs. But fractional designs are much more easily derived for the centroid lattices.

The lattice and centroid designs are not good space fillers. As far as the lattice designs are concerned the degree of the polynomial chosen restricts the observation points. E.g. if one chose to fit a quadratic model one would make observations on pure components and binary mixtures only. A centroid design has only one observation on all the components. Lambrakis's technique of using only observations on q-ary mixtures overcomes this difficulty, but only applies to mixtures of three components or more, requires a large number of observations, leads to some arithmetically tedious equations, although the availability of electronic computers diminishes the importance of the latter.

Draper and Lawrence's designs have a great deal to commend them over those of Lambrakis or Scheffe. One can take into account the existence of bias as well as variance; they are good space fillers and the mathematical technique for dealing with them is well known with softwear readily available. They

fulfil the need expressed by Scheffe for designs with an equally spaced distribution of points. However, at the time of writing designs are available only for three and four components and first and second degree polynomials. These restrictions are not too serious; not many mixtures encountered in practice have more than four components and it has been the writer's experience that where complicated mixtures are encountered it is often due to a willingness to add things to bring about modifications and improvements rather than to remove them. However, some responses do need a third degree polynomial to represent them adequately.

Quenouille has pointed out that one may run into trouble if one of the components is inert. Consider n components which produce no response unless all are present in equal amounts when they produce unit response. The response equation is

$$y = n^n x_1 x_2 \ldots\ldots x_n$$

which has a value of unit at the point $x_1 = x_2 = \ldots = x_n = 1/n$ and zero at all other points.

An inert material is now added so that it forms one half of the total and we have the point

$$(1/2n, 1/2n, \ldots, 1/2n, 1/2)$$

As only 50% active material is present the response will be $\frac{1}{2}$. We have then

$$y = 1 \quad \text{for } x_1 = x_2 = \ldots = x_n = 1/n \qquad x_{n+1} = 0$$

$$y = \tfrac{1}{2} \quad \text{for } x_1 = x_2 = \ldots = x_n = 1/2n \qquad x_{n+1} = \tfrac{1}{2}$$

This implies that the only terms in its equation involve $x_1 x_2 \ldots x_n$ and $x_1 x_2 \ldots x_{n+1}$ and the response equation will be of the form

$$y = Ax_1 x_2 \ldots x_n + Bx_1 x_2 \ldots x_{n+1}$$

Using the first condition

$$1 = A\,n^{-n} + B(0)$$

i.e. $$A = n^{n}$$

Using the second condition

$$\tfrac{1}{2} = A(2n)^{-n} + b(2^{-(n+1)})n^{-n}$$

$$\text{i.e. } B = 2^{n}\,n^{n}(1 - 1/2^{n-1})$$

Whence

$$y = n^{n}x_1x_2 \ldots x_n + 2^{n}\,n^{n}(1-1/2^{n-1})x_1x_2 \ldots x_{n+1}$$

Suppose we now use this equation to predict what happens when all the substances are taken in equal proportions i.e. at $(^1/(n+1), ^1/(n+1), \ldots, ^1/(n+1)$ which is very near to the point $(^1/n, ^1/n, \ldots, ^1/n, 0)$ at which the response was 1. One might reasonably assume that the value of the observed response at this new point would be $n/(n+1)$. However if one inserts the values of the new coordinates in the predictive equation i.e. $x_1 = x_2 = \ldots = x_{n+1} = 1/(n+1)$ one obtains

$$y = \left(\frac{n}{n+1}\right)^{n}\left(1 + \frac{2^{n}}{n+1} - \frac{2}{n+1}\right)$$

For large n the first term tends to e^{-1} while the second becomes very large. Thus, if one takes a 3 component mixture to which a fourth inert component is added the above equation will predict a response of 1.05 while one would have predicted on intuitive grounds that it ought to have been $\frac{3}{4}$!

This illustrates that errors may be aggravated if inert materials which produce no response whatever are included in the design.

Plackett points out that there exist many functions which cannot be conveniently expressed as polynomials e.g. pH of a chemical solution where one is dealing with the negative logarithm of the hydrogen ion activity, or modern theories of drug action.

Plackett and Hewlett (196) derived the following equation for the "effect function" (x) of two drugs z_1 and z_2

$$\exp\{(\alpha_1+\beta_1 \log z_1-x)/\lambda\theta_1\}+\exp\{(\alpha_2+\beta_2 \log z_2-x)/\lambda\theta_2) = 1$$

where $\alpha_1, \alpha_2, \beta_1, \beta_2, \theta_1, \theta_2$ and λ are parameters. Similarly Plackett and Hewlett (1952) obtained the following for a mixture of poisons

$$x = \tfrac{1}{2}(x_1+ x_2)+ \theta \log\left[2 \cosh\left(\frac{x_1-x_2}{2\theta \log_e} \right)\right]$$

where x is itself a function of the dose. Any attempt to fit a polynomial to these expressions would lead to a considerable amount of bias in the sense discussed by Box and Draper.

Scheffe's original work has sparked off a great deal of interest in e xperimental designs for mixtures. The original simplex lattice and simplex centroid designs have the great virtue of simplicity although the generation of the data to enable the response surfaces to be drawn is of such magnitude that a computer is necessary.

The more sophisticated designs more recently produced are of such computational complexity that their use would be unthinkable without a computer. Indeed in the development of Draper and Lawrence's minimum bias and minimum variance designs a computer had to be used to solve the system of non-linear simultaneous equations produced.

The following quotation from Tocher's discussion on the paper by Box and Wilson (1951) is relevant in this context. "Some pragmatists claim that one of the prime purposes of designing experiments is to enable the calculations to be easily performed. It so happens that many of the designs with desirable properties are also easily analysed, but the excuse

for using this property as the aim of designing will soon be swept away with the establishment of modern automatic computing machines as computing aids. These machines would be able to analyse the most complicated experiments in a fraction of an hour."

There are two developments the writer would like to see; a method of fitting a response surface to randomly acquired data analogous to the least squares method of line fitting to unequally spaced data and a computer program which would draw out the response surface on a digital plotter. This is not a trivial problem. Where one has an equation of the form $y = f(x)$ or $y = f(x,z)$ in cartesian coordinates it is easy to arrange for a line for y to be generated for continuous values of x at certain increments of z. But when we have, in the case of mixtures, $y = f(x_1,x_2,1-x_1-x_2)$ for $0 \leqslant x_1, x_2 \leqslant 1$ in triangular coordinates we have to choose a figure for y and find the values of x_1 and x_2 which trace out this contour line. At the very least this entails solving a cubic equation. A suggested method of implementing this is given in Appendix 2

APPENDIX 1.

HADAMARD MATRICES

HADAMARD MATRICES.

A Hadamard matrix of order n is n × n matrix H consisting of 1 and -1 entries such that

$$H H' = n I$$

The inner product of any two distinct rows is zero and the inner product of a row with itself is n. A Hadamard matrix is normalised when the first row and the first column consist of 1's.

$$\begin{vmatrix} 1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ 1 & 1 & -1 & 1 \end{vmatrix} \qquad \begin{vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{vmatrix}$$

NON-NORMAL NORMALISED

A Hadamard matrix of order n (n > 8) is equivalent to a (v,k,λ) symmetrical balanced incomplete block,

$v = n-1$, $k = (n/2)-1$, $\lambda = (n/4)-1$.

Let H be a normalised Hadamard matrix. Delete the first row and the first column and replace all -1 by 0.

This will give a symmetrical balanced incomplete block.

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & 1 & 1 & -1 \\ 1 & -1 & 1 & -1 & -1 & -1 & 1 & 1 \\ 1 & 1 & -1 & 1 & -1 & -1 & -1 & 1 \end{vmatrix} \qquad \begin{vmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 \end{vmatrix}$$

H matrix 8 × 8 B.I.B. from H matrix.

APPENDIX 2.

SUGGESTED ALGORITHM FOR DIRECT PLOTTING OF RESPONSE SURFACES USING AN ON-LINE DIGITAL PLOTTER.

The following is a suggested algorithm for enabling a response surface to be drawn using the digital plotter of a computer.

We will consider the (3,3) simplex lattice design for which the regression equation is

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_{12} x_1 x_2 + \beta_{13} x_1 x_3 + \beta_{23} x_2 x_3 + y_{12} x_1 x_2 (x_1 - x_2) + y_{13} x_1 x_3) + y_{23} x_2 x_3 (x_2 - x_3) + \beta_{123} x_1 x_2 x_3$$

with the conditions $0 \leq x_1, x_2, x_3 \leq 1$

$$x_1 + x_2 + x_3 = 1$$

The problem is that given a contour line y and one coordinate x_1, to find x_2 and x_3. Clearly, we need find only x_2 as x_3 can be found from the relationship that the sum of the coordinates is unity.

Make the substitution $x_3 = 1 - x_1 - x_2$ to obtain a cubic form in x_1 and x_2

$$y = (a_{oo} Z s_{o1} x_2 + a_{o2} x_2^2 + a_{o3} x_2^3) + (a_{1o} + a_{11} x_2 + a_{12} x_2^2) x_1 + (a_{2o} + a_{21} x_2) x_1^2 + a_{3o} x_1^3$$

Z = a
S = x

where

$$a_{oo} = \beta_3$$
$$a_{o1} = \beta_2 - \beta_3 + \beta_{23} - y_{23}$$
$$a_{o2} = -\beta_{23} + 3y_{23}$$
$$a_{93} = -2y_{23}$$
$$a_{1o} = \beta_1 - \beta_2 + \beta_{13} - y_{13}$$
$$a_{11} = \beta_{12} - \beta_{13} - \beta_{23} + 2y_{13} + 2y_{23} + \beta_{123}$$
$$a_{12} = -y_{12} - y_{13} - 3y_{23} - \beta_{123}$$
$$a_{2o} = -\beta_{13} + 3y_{13}$$
$$a_{21} = y_{12} - 3y_{13} - y_{23} - \beta_{123}$$
$$a_{3o} = -2y_{13}$$

All these coefficients are known, hence all the a's may be calculated. x_2 is also known and the problem then is to solve the cubic equation in x_1

$$d_0 + d_1x_1 + d_2x_1^2 + d_3x_1^3 = 0$$

where

$$d_0 = (a_{00} + a_{01}x_2 + a_{02}x_2^2 + a_{03}x_2^3) - y$$

$$d_1 = (a_{10} + a_{11}x_2 + a_{12}x_2^2)$$

$$d_2 = (a_{20} + a_{21}x_2)$$

$$d_3 = a_{30}$$

for real roots such as $0 \leqslant x_1 \leqslant 1$. x_2 is already specified and $x_3 = 1 - x_1 - x_2$.

The cubic in x_1 may be further simplified by making the substitution $x_1 = x_1 - d_2/3d_3$ which yields the cubic

$$d_3x_1^3 + \left(\frac{d_2^2}{3d_3} - \frac{2d_2^2}{3d_3} + d_1\right)x_1 + \left(\frac{2d_2^3}{27d_3^2} - \frac{d_1d_2}{3d_3} + d_0\right) = 0$$

in which the quadratic term has been eliminated. This is probably not necessary as computers usually have subroutines for the solution of polynomials of degree higher than is envisaged in this work.

Now, digital plotters usually have procedures by which it is possible to mark points specified in cartesian coordinates and it is probably easier to make the simple conversion from triangular coordinates (x_1, x_2, x_3) to cartesian coordinates (z_1, z_2) for the purposes of plotting using the formulae

$$z_1 = (-3x_1 - x_2\sqrt{3}+1)/3;\ z_2 = (3x_1 - x_2\sqrt{3}+1)/3;\ z_3 = (2x_2\sqrt{3}+1).$$

As regards the values for the response contours that are to be drawn; suitable values at equal intervals will be suggested by the observed responses and experience suggests

that about 10-15 lines on the simplex are usually adequate.

If now the simplex is scanned at some suitable interval (0.02 has been found to be adequately small and in some cases 0.05 will suffice) and the cubic equation solved for real roots at each point for the various values of the response contours (y's) chosen, the coordinates of the points may be stored in an array. Use may then be made of the procedure which exists for digital plotters whereby a smooth curve is drawn through a given set of points by fitting a cubic polynomial between (x_i,y_i) and (x_{i+1},y_{i+1}) using the four coordinates $(x_{i-1},y_{i-1}),(x_{i+1},y_{i+1})$ and (x_{i+2},y_{i+2}).

APPENDIX 3.

AN ALTERNATIVE METHOD OF CALCULATING THE VARIABLE COEFFICIENTS WHEN PROCESS VARIABLES ARE PRESENT

The reading of Scheffe's paper (1963) leads to the following method of calculating the β coefficients when process variables are present. The results are identical with those given on page 41 which seem more natural in their derivation. An illustration of the calculation of the β_{13} coefficient is given so as to avoid any confusion on this point. Scheffe's notation will be used which is self-explanatory apart from such coefficients as $\beta_{23}^{AB,ij}$ which means the coefficient β_{23} with A at the i^{th} level and B at the j^{th} level more usually seen as [(ab+(1))-(a+b)].

The responses from Block 1, Block 3 and Block 5 are taken from Table 3.

Block 1		Block 3		Block 5	
$y_{1,..}$	= 100	$y_{3,..}$	= 85	$y_{13,..}$	= 90
$y_{1,i.}$	= 108	$y_{3,i.}$	= 98	$y_{13,i.}$	= 120
$y_{1,.j}$	= 186	$y_{3,.j}$	= 129	$y_{13,j}$	= 130
$y_{1,ij}$	= 207	$y_{3,ij}$	= 157	$y_{13,ij}$	= 190

Using the formulae for calculating the coefficients of the special cubic model we have for $S_r = 13$,

$$\beta_{..} = 4\times90 - 2\times100 - 2\times85 = -10$$
$$\beta_{i.} = 4\times120 - 2\times108 - 2\times98 = 68$$
$$\beta_{j.} = 4\times130 - 2\times186 - 2\times129 = -110$$
$$\beta_{ij} = 4\times90 - 2\times207 - 2\times157 = 32$$

The various effects and interactions are therefore,

$$\beta_{13}^{A,i} = \tfrac{1}{4}((68+32) - (-110-10)) = 55$$

$$\beta_{13}^{B,j} = \tfrac{1}{4}((-110+32) - (68-10)) = -34$$

$$\beta_{13}^{AB,ij} = \tfrac{1}{4}((-10+32) - (68-110)) = 16$$

$$\beta_{13}^{o} = \tfrac{1}{4}(-10+68-110+32) = -5$$

and taking these figures

$$\beta_{13,ij} = -5+55A - 34B + 16AB$$

which is identical with the β_{13} coefficient given on page 41.

APPENDIX 4 .

When equal numbers of observations are made at each lattice point $\mathrm{Var}(\tilde{y}) = \sigma^2 z/r$ where r is the number of observations at each point and z is the sum of the squares of the coefficients given on page 25

These tables give values of z for the quadratic, special cubic, cubic and quartic models for the intervals 1.0(0.02)0.32.

CO-ORDINATES			QUAD.	SP. CUBIC	CUBIC	QUARTIC
1.00	0.00	0.00	1.0000	1.0000	1.0000	1.0000
0.98	0.00	0.02	0.8916	0.8916	0.8349	0.8488
0.96	0.00	0.04	0.8050	0.8050	0.7604	0.9398
	0.02	0.02	0.7926	0.7933	0.6974	0.7354
0.94	0.00	0.06	0.7379	0.7379	0.7507	1.1340
	0.02	0.04	0.7143	0.7143	0.6409	0.8274
0.92	0.00	0.08	0.6884	0.6884	0.7846	1.3410
	0.02	0.06	0.6546	0.6527	0.6418	1.0023
	0.04	0.04	0.6433	0.6414	0.5948	0.8973
0.90	0.00	0.10	0.6544	0.6544	0.8447	1.5072
	0.02	0.08	0.6115	0.6067	0.6803	1.1814
	0.04	0.06	0.5900	0.5846	0.6000	1.0396
0.88	0.00	0.12	0.6340	0.6340	0.9171	1.6063
	0.02	0.10	0.5830	0.5745	0.7405	1.3193
	0.04	0.08	0.5525	0.5420	0.6381	1.1843
	0.06	0.06	0.5423	0.5317	0.6046	1.1455
0.86	0.00	0.14	0.6255	0.6255	0.9910	1.6317
	0.02	0.12	0.5673	0.5544	0.8098	1.3950
	0.04	0.10	0.5287	0.5120	0.6946	1.2921
	0.06	0.08	0.5094	0.4924	0.6387	1.2568
0.84	0.00	0.16	0.6271	0.6271	1.0582	1.5897
	0.02	0.14	0.5627	0.5447	0.8784	1.4048
	0.04	0.12	0.5169	0.4931	0.7578	1.3453
	0.06	0.10	0.4896	0.4651	0.6887	1.3380
	0.08	0.08	0.4805	0.4562	0.6661	1.3402
0.82	0.00	0.18	0.6373	0.6373	1.1132	1.4945
	0.02	0.16	0.5675	0.5439	0.9391	1.3561
	0.04	0.14	0.5155	0.4837	0.8191	1.3416
	0.06	0.12	0.4810	0.4482	0.7440	1.3732
	0.08	0.10	0.4638	0.4318	0.7080	1.4011
0.80	0.00	0.20	0.6544	0.6544	1.1523	1.3638
	0.02	0.18	0.5801	0.5507	0.9871	1.2628
	0.04	0.16	0.5228	0.4824	0.8719	1.2884
	0.06	0.14	0.4820	0.4402	0.7968	1.3604
	0.08	0.12	0.4577	0.4175	0.7545	1.4246
	0.10	0.10	0.4496	0.4104	0.7409	1.4494

CO-ORDINATES			QUAD.	SP. CUBIC	CUBIC	QUARTIC
0·78	0·00	0·22	0·6771	0·6771	1·1737	1·2163
	0·02	0·20	0·5991	0·5636	1·0194	1·1416
	0·04	0·18	0·5372	0·4879	0·9122	1·1986
	0·06	0·16	0·4911	0·4399	0·8413	1·3065
	0·08	0·14	0·4606	0·4120	0·7985	1·4084
	0·10	0·12	0·4453	0·3994	0·7784	1·4680
0·76	0·00	0·24	0·7041	0·7041	1·1770	1·0690
	0·02	0·22	0·6230	0·5814	1·0347	1·0092
	0·04	0·20	0·5574	0·4989	0·9375	1·0874
	0·06	0·18	0·5068	0·4458	0·8738	1·2229
	0·08	0·16	0·4709	0·4138	0·8346	1·3579
	0·10	0·14	0·4494	0·3970	0·8137	1·4536
	0·12	0·12	0·4423	0·3918	0·8071	1·4880
0·74	0·00	0·26	0·7340	0·7340	1·1631	0·9361
	0·02	0·24	0·6507	0·6029	1·0330	0·8808
	0·04	0·22	0·5820	0·5142	0·9468	0·9695
	0·06	0·20	0·5277	0·4568	0·8923	1·1228
	0·08	0·18	0·4872	0·4216	0·8599	1·2829
	0·10	0·16	0·4604	0·4019	0·8422	1·4102
	0·12	0·14	0·4471	0·3932	0·8346	1·4799
0·72	0·00	0·28	0·7658	0·7658	1·1337	0·8280
	0·02	0·26	0·6808	0·6270	1·0153	0·7685
	0·04	0·24	0·6098	0·5328	0·9405	0·8581
	0·06	0·22	0·5524	0·4718	0·8963	1·0190
	0·08	0·20	0·5082	0·4342	0·8724	1·1948
	0·10	0·18	0·4769	0·4127	0·8610	1·3459
	0·12	0·16	0·4582	0·4019	0·8564	1·4463
	0·14	0·14	0·4520	0·3986	0·8553	1·4813
0·70	0·00	0·30	0·7984	0·7984	1·0912	0·7511
	0·02	0·28	0·7122	0·6526	0·9838	0·6808
	0·04	0·26	0·6396	0·5537	0·9198	0·7631
	0·06	0·24	0·5799	0·4896	0·8862	0·9227
	0·08	0·22	0·5327	0·4504	0·8720	1·1045
	0·10	0·20	0·4976	0·4280	0·8687	1·2701
	0·12	0·18	0·4744	0·4165	0·8701	1·3935
	0·14	0·16	0·4628	0·4117	0·8718	1·4590

CO-ORDINATES.			QUAD.	SP. CUBIC	CUBIC	QUARTIC
0.68	0.00	0.32	0.8308	0.8308	1.0389	0.7077
	0.02	0.30	0.7441	0.6789	0.9409	0.6222
	0.04	0.28	0.6703	0.5758	0.8869	0.6914
	0.06	0.26	0.6089	0.5094	0.8634	0.8421
	0.08	0.24	0.5594	0.4693	0.8591	1.0214
	0.10	0.22	0.5213	0.4468	0.8651	1.1922
	0.12	0.20	0.4943	0.4355	0.8742	1.3296
	0.14	0.18	0.4782	0.4306	0.8817	1.4178
	0.16	0.16	0.4729	0.4293	0.8845	1.4481
0.66	0.00	0.34	0.8621	0.8621	0.9798	0.6965
	0.02	0.32	0.7754	0.7049	0.8896	0.5938
	0.04	0.30	0.7011	0.5984	0.8443	0.6462
	0.06	0.28	0.6386	0.5302	0.8301	0.7826
	0.08	0.26	0.5874	0.4897	0.8354	0.9522
	0.10	0.24	0.5469	0.4679	0.8507	1.1199
	0.12	0.22	0.5169	0.4578	0.8687	1.2623
	0.14	0.20	0.4971	0.4540	0.8837	1.3643
	0.16	0.18	0.4872	0.4530	0.8921	1.4173
0.64	0.00	0.36	0.8916	0.8916	0.9175	0.7134
	0.02	0.34	0.8053	0.7300	0.8332	0.5934
	0.04	0.32	0.7309	0.6207	0.7949	0.6274
	0.06	0.30	0.6679	0.5513	0.7887	0.7463
	0.08	0.28	0.6156	0.5109	0.8027	0.9013
	0.10	0.26	0.5735	0.4903	0.8271	1.0592
	0.12	0.24	0.5411	0.4820	0.8541	1.1983
	0.14	0.22	0.5183	0.4803	0.8776	1.3053
	0.16	0.20	0.5047	0.4809	0.8934	1.3721
	0.18	0.18	0.5001	0.4814	0.8990	1.3948
0.62	0.00	0.38	0.9186	0.9186	0.8553	0.7518
	0.02	0.36	0.8331	0.7534	0.7750	0.6163
	0.04	0.34	0.7591	0.6420	0.7420	0.6325
	0.06	0.32	0.6960	0.5718	0.7421	0.7326
	0.08	0.30	0.6432	0.5320	0.7636	0.8700
	0.10	0.28	0.6000	0.5130	0.7962	1.0133
	0.12	0.26	0.5659	0.5072	0.8319	1.1428
	0.14	0.24	0.5408	0.5083	0.8641	1.2468
	0.16	0.22	0.5241	0.5117	0.8883	1.3187
	0.18	0.20	0.5158	0.5141	0.9011	1.3553
0.60	0.00	0.40	0.9424	0.9424	0.7965	0.8040
	0.02	0.38	0.8581	0.7746	0.7182	0.6561
	0.04	0.36	0.7849	0.6617	0.6885	0.6564
	0.06	0.34	0.7223	0.5912	0.6933	0.7386

CO-ORDINATES			QUAD.	SP. CUBIC	CUBIC	QUARTIC
0.60	0.08	0.32	0.6694	0.5523	0.7207	0.8574
	0.10	0.30	0.6256	0.5353	0.7603	0.9834
	0.12	0.28	0.5905	0.5324	0.8038	1.0987
	0.14	0.26	0.5636	0.5369	0.8444	1.1934
	0.16	0.24	0.5446	0.5438	0.8771	1.2626
	0.18	0.22	0.5333	0.5494	0.8982	1.3045
	0.20	0.20	0.5296	0.5515	0.9054	1.3185
0.58	0.00	0.42	0.9626	0.9626	0.7440	0.8614
	0.02	0.40	0.8797	0.7931	0.6657	0.7052
	0.04	0.38	0.8078	0.6793	0.6375	0.6930
	0.06	0.36	0.7460	0.6089	0.6451	0.7595
	0.08	0.34	0.6935	0.5711	0.6767	0.8607
	0.10	0.32	0.6496	0.5564	0.7218	0.9686
	0.12	0.30	0.6140	0.5566	0.7719	1.0671
	0.14	0.28	0.5860	0.5650	0.8202	1.1481
	0.16	0.26	0.5654	0.5760	0.8610	1.2084
	0.18	0.24	0.5517	0.5856	0.8904	1.2480
	0.20	0.22	0.5450	0.5911	0.9058	1.2674
0.56	0.00	0.44	0.9787	0.9787	0.7004	0.9160
	0.02	0.42	0.8976	0.8083	0.6203	0.7559
	0.04	0.40	0.8272	0.6942	0.5916	0.7354
	0.06	0.38	0.7665	0.6244	0.6003	0.7897
	0.08	0.36	0.7149	0.5880	0.6343	0.8756
	0.10	0.34	0.6715	0.5757	0.6832	0.9665
	0.12	0.32	0.6358	0.5792	0.7386	1.0475
	0.14	0.30	0.6072	0.5916	0.7932	1.1122
	0.16	0.28	0.5855	0.6071	0.8413	1.1594
	0.18	0.26	0.5702	0.6213	0.8788	1.1905
	0.20	0.24	0.5611	0.6311	0.9025	1.2078
	0.22	0.22	0.5581	0.6346	0.9106	1.2134
0.54	0.00	0.46	0.9905	0.9905	0.6677	0.9607
	0.02	0.44	0.9114	0.8201	0.5842	0.8011
	0.04	0.42	0.8427	0.7062	0.5533	0.7768
	0.06	0.40	0.7836	0.6372	0.5613	0.8231
	0.08	0.38	0.7331	0.6025	0.5959	0.8972
	0.10	0.36	0.6905	0.5926	0.6470	0.9734
	0.12	0.34	0.6552	0.5994	0.7059	1.0380
	0.14	0.32	0.6266	0.6159	0.7654	1.0859
	0.16	0.30	0.6043	0.6361	0.8198	1.1175
	0.18	0.28	0.5879	0.6553	0.8644	1.1359
	0.20	0.26	0.5771	0.6701	0.8960	1.1451
	0.22	0.24	0.5718	0.6781	0.9123	1.1487

CO-ORDINATES			QUAD,	SP. CUBIC	CUBIC	QUARTIC
0·52	0·00	0·48	0·9976	0·9976	0·6475	0·9899
	0·02	0·46	0·9207	0·8280	0·5591	0·8348
	0·04	0·44	0·8540	0·7150	0·5246	0·8112
	0·06	0·42	0·7966	0·6472	0·5301	0·8539
	0·08	0·40	0·7476	0·6141	0·5637	0·9205
	0·10	0·38	0·7063	0·6067	0·6152	0·9855
	0·12	0·36	0·6718	0·6168	0·6760	1·0361
	0·14	0·34	0·6436	0·6372	0·7388	1·0682
	0·16	0·32	0·6213	0·6621	0·7979	1·0835
	0·18	0·30	0·6043	0·6865	0·8486	1·0867
	0·20	0·28	0·5924	0·7067	0·8872	1·0836
	0·22	0·26	0·5854	0·7198	0·9114	1·0794
	0·24	0·24	0·5831	0·7244	0·9196	1·0776
0·50	0·00	0·50	1·0000	1·0000	0·6406	1·0000
	0·02	0·48	0·9254	0·8321	0·5463	0·8528
	0·04	0·46	0·8609	0·7204	0·5067	0·8338
	0·06	0·44	0·8055	0·6539	0·5084	0·8774
	0·08	0·42	0·7583	0·6227	0·5393	0·9409
	0·10	0·40	0·7184	0·6176	0·5896	0·9988
	0·12	0·38	0·6851	0·6307	0·6506	1·0387
	0·14	0·36	0·6578	0·6550	0·7152	1·0573
	0·16	0·34	0·6358	0·6844	0·7775	1·0573
	0·18	0·32	0·6188	0·7139	0·8327	1·0449
	0·20	0·30	0·6064	0·7396	0·8773	1·0271
	0·22	0·28	0·5983	0·7584	0·9085	1·0108
	0·24	0·26	0·5943	0·7683	0·9245	1·0013
0·48	0·00	0·52	0·9976	0·9976	0·6475	0·9899
	0·02	0·50	0·9254	0·8321	0·5463	0·8528
	0·04	0·48	0·8632	0·7222	0·5007	0·8417
	0·06	0·46	0·8100	0·6573	0·4972	0·8901
	0·08	0·44	0·7648	0·6279	0·5242	0·9547
	0·10	0·42	0·7267	0·6250	0·5717	1·0097
	0·12	0·40	0·6949	0·6409	0·6314	1·0429
	0·14	0·38	0·6687	0·6687	0·6960	1·0514
	0·16	0·36	0·6476	0·7024	0·7599	1·0384
	0·18	0·34	0·6310	0·7368	0·8182	1·0111
	0·20	0·32	0·6185	0·7679	0·8673	0·9780
	0·22	0·30	0·6098	0·7924	0·9043	0·9474
	0·24	0·28	0·6047	0·8081	0·9273	0·9260
	0·26	0·26	0·6030	0·8134	0·9351	0·9183

CO-ORDINATES			QUAD.	SP.CUBIC	CUBIC	QUARTIC
0.46	0.00	0.54	0.9905	0.9905	0.6677	0.9607
	0.02	0.52	0.9207	0.8280	0.5591	0.8348
	0.04	0.50	0.8609	0.7204	0.5067	0.8338
	0.06	0.48	0.8100	0.6573	0.4972	0.8901
	0.08	0.46	0.7670	0.6296	0.5191	0.9595
	0.10	0.44	0.7308	0.6288	0.5626	1.0159
	0.12	0.42	0.7008	0.6472	0.6193	1.0465
	0.14	0.40	0.6761	0.6781	0.6825	1.0486
	0.16	0.38	0.6562	0.7156	0.7464	1.0258
	0.18	0.36	0.6404	0.7545	0.8062	0.9859
	0.20	0.34	0.6284	0.7908	0.8583	0.9384
	0.22	0.32	0.6197	0.8210	0.8998	0.8927
	0.24	0.30	0.6140	0.8425	0.9285	0.8571
	0.26	0.28	0.6112	0.8536	0.9433	0.8378
0.44	0.00	0.56	0.9787	0.9787	0.7004	0.9160
	0.02	0.54	0.9114	0.8201	0.5842	0.8011
	0.04	0.52	0.8540	0.7150	0.5246	0.8112
	0.06	0.50	0.8055	0.6539	0.5084	0.8774
	0.08	0.48	0.7648	0.6279	0.5242	0.9547
	0.10	0.46	0.7308	0.6288	0.5626	1.0159
	0.12	0.44	0.7028	0.6493	0.6153	1.0479
	0.14	0.42	0.6799	0.6828	0.6756	1.0475
	0.16	0.40	0.6615	0.7236	0.7379	1.0187
	0.18	0.38	0.6469	0.7666	0.7976	0.9692
	0.20	0.36	0.6356	0.8076	0.8511	0.9093
	0.22	0.34	0.6273	0.8431	0.8956	0.8495
	0.24	0.32	0.6216	0.8704	0.9288	0.7991
	0.26	0.30	0.6183	0.8875	0.9493	0.7657
0.42	0.00	0.58	0.9626	0.9626	0.7440	0.8614
	0.02	0.56	0.8976	0.8083	0.6203	0.7559
	0.04	0.54	0.8427	0.7062	0.5533	0.7768
	0.06	0.52	0.7966	0.6472	0.5301	0.8539
	0.08	0.50	0.7583	0.6227	0.5393	0.9409
	0.10	0.48	0.7267	0.6250	0.5717	1.0097
	0.12	0.46	0.7008	0.6472	0.6193	1.0465
	0.14	0.44	0.6799	0.6828	0.6756	1.0475
	0.16	0.42	0.6632	0.7263	0.7350	1.0163
	0.18	0.40	0.6502	0.7727	0.7932	0.9609
	0.20	0.38	0.6401	0.8179	0.8466	0.8917
	0.22	0.36	0.6326	0.8582	0.8924	0.8196
	0.24	0.34	0.6273	0.8909	0.9285	0.7552
	0.26	0.32	0.6239	0.9139	0.9534	0.7071
	0.28	0.30	0.6223	0.9257	0.9661	0.6815

CO-ORDINATES			QUAD.	SP.CUBIC	CUBIC	QUARTIC
0·40	0·00	0·60	0·9424	0·9424	0·7965	0·8040
	0·02	0·58	0·8797	0·7931	0·6657	0·7052
	0·04	0·56	0·8272	0·6942	0·5916	0·7354
	0·06	0·54	0·7836	0·6372	0·5613	0·8231
	0·08	0·52	0·7476	0·6141	0·5637	0·9205
	0·10	0·50	0·7184	0·6176	0·5896	0·9988
	0·12	0·48	0·6949	0·6409	0·6314	1·0429
	0·14	0·46	0·6761	0·6781	0·6825	1·0486
	0·16	0·44	0·6615	0·7236	0·7379	1·0187
	0·18	0·42	0·6502	0·7727	0·7932	0·9609
	0·20	0·40	0·6416	0·8213	0·8450	0·8857
	0·22	0·38	0·6353	0·8658	0·8907	0·8044
	0·24	0·36	0·6308	0·9035	0·9282	0·7280
	0·26	0·34	0·6278	0·9319	0·9560	0·6659
	0·28	0·32	0·6261	0·9496	0·9730	0·6256
0·38	0·00	0·62	0·9186	0·9186	0·8553	0·7518
	0·02	0·60	0·8581	0·7746	0·7182	0·6561
	0·04	0·58	0·8078	0·6793	0·6375	0·6930
	0·06	0·56	0·7665	0·6244	0·6003	0·7897
	0·08	0·54	0·7331	0·6025	0·5959	0·8972
	0·10	0·52	0·7063	0·6067	0·6152	0·9855
	0·12	0·50	0·6851	0·6307	0·6506	1·0387
	0·14	0·48	0·6687	0·6687	0·6960	1·0514
	0·16	0·46	0·6562	0·7156	0·7464	1·0258
	0·18	0·44	0·6469	0·7666	0·7976	0·9692
	0·20	0·42	0·6401	0·8179	0·8466	0·8917
	0·22	0·40	0·6353	0·8658	0·8907	0·8044
	0·24	0·38	0·6320	0·9077	0·9280	0·7187
	0·26	0·36	0·6298	0·9411	0·9572	0·6447
	0·28	0·34	0·6286	0·9643	0·9771	0·5905
	0·30	0·32	0·6280	0·9762	0·9873	0·5619
0·36	0·00	0·64	0·8916	0·8916	0·9175	0·7134
	0·02	0·62	0·8331	0·7534	0·7750	0·6163
	0·04	0·60	0·7849	0·6617	0·6885	0·6564
	0·06	0·58	0·7460	0·6089	0·6451	0·7595
	0·08	0·56	0·7149	0·5880	0·6343	0·8756
	0·10	0·54	0·6905	0·5926	0·6470	0·9734
	0·12	0·52	0·6718	0·6168	0·6760	1·0361
	0·14	0·50	0·6578	0·6550	0·7152	1·0573
	0·16	0·48	0·6476	0·7024	0·7599	1·0384
	0·18	0·46	0·6404	0·7545	0·8062	0·9859
	0·20	0·44	0·6356	0·8076	0·8511	0·9093
	0·22	0·42	0·6326	0·8582	0·8924	0·8196
	0·24	0·40	0·6308	0·9035	0·9282	0·7280
	0·26	0·38	0·6298	0·9411	0·9572	0·6447
	0·28	0·36	0·6294	0·9693	0·9785	0·5785
	0·30	0·34	0·6292	0·9867	0·9915	0·5361

CO-ORDINATES			QUAD$_3$	SP. CUBIC	CUBIC	QUARTIC
0·34	0·00	0·66	0·8621	0·8621	0·9798	0·6965
	0·02	0·64	0·8053	0·7300	0·8332	0·5934
	0·04	0·62	0·7591	0·6420	0·7420	0·6325
	0·06	0·60	0·7223	0·5912	0·6933	0·7386
	0·08	0·58	0·6935	0·5711	0·6767	0·8607
	0·10	0·56	0·6715	0·5757	0·6832	0·9665
	0·12	0·54	0·6552	0·5994	0·7059	1·0380
	0·14	0·52	0·6436	0·6372	0·7388	1·0682
	0·16	0·50	0·6358	0·6844	0·7775	1·0573
	0·18	0·48	0·6310	0·7368	0·8182	1·0111
	0·20	0·46	0·6284	0·7908	0·8583	0·9384
	0·22	0·44	0·6273	0·8431	0·8956	0·8495
	0·24	0·42	0·6273	0·8909	0·9285	0·7552
	0·26	0·40	0·6278	0·9319	0·9560	0·6659
	0·28	0·38	0·6286	0·9643	0·9771	0·5905
	0·30	0·36	0·6292	0·9867	0·9915	0·5361
	0·32	0·34	0·6296	0·9981	0·9988	0·5076
0·32	0·00	0·68	0·8308	0·8308	1·0389	0·7077
	0·02	0·66	0·7754	0·7049	0·8896	0·5938
	0·04	0·64	0·7309	0·6207	0·7949	0·6274
	0·06	0·62	0·6960	0·5718	0·7421	0·7326
	0·08	0·60	0·6694	0·5523	0·7207	0·8574
	0·10	0·58	0·6496	0·5564	0·7218	0·9686
	0·12	0·56	0·6358	0·5792	0·7386	1·0475
	0·14	0·54	0·6266	0·6159	0·7654	1·0859
	0·16	0·52	0·6213	0·6621	0·7979	1·0835
	0·18	0·50	0·6188	0·7139	0·8327	1·0449
	0·20	0·48	0·6185	0·7679	0·8673	0·9780
	0·22	0·46	0·6197	0·8210	0·8998	0·8927
	0·24	0·44	0·6216	0·8704	0·9288	0·7991
	0·26	0·42	0·6239	0·9139	0·9534	0·7071
	0·28	0·40	0·6261	0·9496	0·9730	0·6256
	0·30	0·38	0·6280	0·9762	0·9873	0·5619
	0·32	0·36	0·6291	0·9926	0·9959	0·5215
0·30	0·00	0·70	0·7984	0·7984	1·0912	0·7511
	0·02	0·68	0·7441	0·6789	0·9409	0·6222
	0·04	0·66	0·7011	0·5984	0·8443	0·6462
	0·06	0·64	0·6679	0·5513	0·7887	0·7463
	0·08	0·62	0·6432	0·5320	0·7636	0·8700
	0·10	0·60	0·6256	0·5353	0·7603	0·9834
	0·12	0·58	0·6140	0·5566	0·7719	1·0671
	0·14	0·56	0·6072	0·5916	0·7932	1·1122
	0·16	0·54	0·6043	0·6361	0·8198	1·1175
	0·18	0·52	0·6043	0·6865	0·8486	1·0867
	0·20	0·50	0·6064	0·7396	0·8773	1·0271
	0·22	0·48	0·6098	0·7924	0·9043	0·9474
	0·24	0·46	0·6140	0·8425	0·9285	0·8571
	0·26	0·44	0·6183	0·8875	0·9493	0·7657
	0·28	0·42	0·6223	0·9257	0·9661	0·6815
	0·30	0·40	0·6256	0·9556	0·9788	0·6117
	0·32	0·38	0·6280	0·9762	0·9873	0·5619
	0·34	0·36	0·6292	0·9867	0·9915	0·5361

REFERENCES.

BOX,G.E.P., DRAPER,N.R., (1959) "A Basis for the Selection of a Response Surface Design", J.A.S.A.,54,622.

BOX,G.E.P., HUNTER,J.S.,(1957) "Multifactor Experimental Designs for Exploring Response Surfaces", Ann.Math,Stat.,28,195.

BOX,G.E.P., WILSON,K.B.,(1951) "On the Experimental Attainment of Optimum Conditions", J.R.S.S., B13(1),1.

DRAPER,N.R.,LAWRENCE,W.,(1965) "Mixture Designs for Three Factors", J.R.S.S.,B27(3),450.

DRAPER,N.R.,LAWRENCE,W.,(1965) "Mixture Designs for Four Factors, J.R.S.S.,B27(3),473.

DOEHLERT,D.H.,(1970) "Uniform Shell Designs", J.R.S.S.,C19(3),231.

DOEHLERT,D.H.,KLEE,V.L.,(1970) "Experimental Designs Through Level Reduction of the d-dimensional Cuboctahedron" Unpublished Work.

DUNN,O.J.,(1959), "Confidence Intervals for the Means of Dependent Normally Distributed Variables," J.A.S.A.,54,613.

idem (1968) "A Note on Confidence Bands for a Regression Line Over a Finite Range", J.A.S.A.,63,1028.

EPERSON,D.B.,(1967) "Triangular Numbers", Math.Gazette,51(377),242.

FEDERIGHI,E.T.,(1959) "Extended Tables of the Percentage Points of Student's t-distribution" J.A.S.A.,54,683.

FISHER,R.A.,(1925), "Expansion of Student's Integral in Powers of n^{-1}, Metron, 5,109.

GORMAN,J.W., HINMAN,J.E.,(1962), "Simplex Lattice Designs for Multicomponent Systems", Technometrics,4(4),463.

HARRINGTON,E.C.,pp.175-6 "Evoluntionary Operation" by G.E.P.Box and N.R.Draper, J.W.Wiley,1969.

LAMBRAKIS,D.P.,(1968), "Experiments with p-Component Mixtures", J.R.S.S.,B30(1),137.

idem (1968), "Experiments with Mixtures: A Generalisation of the Simplex Lattice Design", J.R.S.S.,B30(1),123.

LOEWE,S.,(1968), "Die Quantitativen Probleme der Pharmakologie", Ergeb, Physiol., 27,47.

LUI,C.L.,"Introduction to Combinatorial Analysis" McGraw Hill, 1968.

MACHT,D.I.,(1929), "Pharmocological Synergism of Stereoisomers", Proc.Nat.Acad.Sci.Wash.,195,63.

PLACKETT,R.L.,HEWLETT,P.S.,(1952), "Quantal Responses to Mixtures of Poisons", J.R.S.S.,B.14(2),141.

idem (1967) "A Comparison of two Approaches to the Construction of Models for Quantal Responses to Mixtures of Drugs", Biometrics, 23(1),26.

QUENOUILLE,M.H.,(1959), "Experiments with Mixtures", J.R.S.S., B.21(1),201.

SCHEFFE,H.,(1958), "Experiments with Mixtures", J.R.S.S., B.20(2),344.

idem (1961), "Reply to Mr.(Sic) Quenouille's comments about my Paper on Mixtures", J.R.S.S.,B.23(1),171.

idem (1963) "The Simplex Centroid Design for Experiments with Mixtures", J.R.S.S.,B.25(2),235.

SNEE,R.D., (1971) "The Design and Analysis of Mixture Experiments J.of Quality Technology, 3(4), October.

NOTE 1. The paper by Gorman and Hinman contains three errors. Page 486 under Special Cubic Model, the equation for b_i should read

$$b_i = (x_i/2)\left(6x_i^2 - 2x_i + 1 - 3\sum_{i \leqslant 1 \leqslant 3} x_1^2\right)$$

and under Full Cubic Model the second term in Var(η) should be eliminated.

On page 478 a negative sign is missing from the tenth term of the quartic equation.

NOTE 2. The paper by Scheffe (1958) contains two errors.

Page 349, in formula 4.8, the coefficient should 413 should be -12 and the coefficient -3 should be +3.

NOTE 3. The papers by Loewe and Macht have not been consulted by the writer. They are quoted as the original sources of the use of "synergistic" and "isobols".

NOTE 4. The paper by Dunn (1959) contains an error in the first column of the table on page 618. For $S_i\sqrt{10}$ read $S_i/\sqrt{10}$.

NOTE 5. The paper by Box and Draper (1959) contains an error in equation 5. It is correctly stated in equation 7.

NOTE 6. The paper by Box and Draper (1963) contains an error in the third equation of Appendix 1. For $x_1 y_1$ read $x_1' y_1$

NOTE 7. The book by Box and Draper, "Evolutionary Operation" contains three misprints in the title to Appendix 1.

NOTE 8. The paper by Snee has not been consulted due to the difficulty experienced in obtaining a copy. This reference is given for the sake of completeness.