STATISTICAL APPROACHES TO CERTAIN PROBLEMS IN GEOPHYSICS

by

STEPHEN MILTON SIMPSON, JR.

B.S., YALE UNIVERSITY

(1950)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

(1953)

Signature of Author . . . .
Department of Geology and Geophysics
August 14, 1953

Certified by . . . .
Thesis Supervisor

. . . .
Chairman, Departmental Committee on Graduate Students

ABSTRACT

STATISTICAL APPROACHES TO CERTAIN PROBLEMS IN GEOPHYSICS
by
Stephen Milton Simpson, Jr.

Submitted to the Department of Geology and Geophysics on August 14, 1953, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Several specific problems in seismic and gravitational data interpretation are considered from the statistical viewpoint. Least squares techniques are applied to the two types of interpretation, and, for seismic records, other approaches are discussed.

The fitting of an nth order polynomial in x and y to gravity data by the method of least squares is investigated as a method for approximating regional gravitational anomalies. The normal equations for the general case are derived and simplification considered. It is shown that, with a symmetrical rectangular distribution of gravity readings, each set of these equations breaks up into smaller subsets. The resulting simplification brings fairly high order polynomials into the range of practical computation. For a particular gridwork the polynomial coefficients may be expressed explicitly as linear combinations of the right hand members of the normal equations. Once this is done, the least squares fitting of any data taken over such a gridwork may be effected relatively easily. The explicit expressions for the coefficients are derived for a square gridwork of 121 points and for polynomials of order 2, 3, and 4. A set of actual gravity readings is analysed in this fashion. The gravity residuals are determined and contoured. The comparison of these contours with each other (for various order polynomials) and with contours derived by a standard, much more involved process, is favorable. This consistency, despite certain detrimental features of the data used indicates that the method may deserve to find practical application as a routine, first step, gravity reduction procedure. The problem is pursued with regard to different gridworks, and a table is derived which contains, in effect, the normal equations for representative grids up to a size containing 2601 points, and for polynomials through order four.

As an approach to the understanding of linear operators, as they apply to the analysis of seismic records, a simple form of linear operator is studied. For this form of operator, the so-called "cosine operator", certain properties are derived in the general case, and interpreted geometrically. These include relationships between the exact form of cosine operator chosen, the correlation properties of the series which the operator is to predict, the individual errors of prediction, and the sums of squared errors of prediction. The results are applied to two classes of time series in connection with spectrum analysis, and, for one class, filter characteristics are computed for a specific cosine operator.

An iterative method for determining least squares fits of linear operators to multiple time series is discussed geometrically. An argument is presented, based on the geometry of the two term operator, to show that, in the case of near singularity where many solutions will almost satisfy the least squares criterion, the exact solution is necessary for the purposes under consideration.

Several interpretive procedures are devised for finding information from seismic records. The first deals with discriminating an unknown velocity in a two velocity system. An adaption is made for detecting reflections, and practical example are given of the two uses. The second employs a form of testing phase, between seismic traces and their predictions by linear operators, to determine reflection times, and is illustrated with an example. The third combines the concept of ensemble averages with linear operators to determine step-out times of reflections. The last procedure suggests a special experimental arrangement, coupled with a certain type of correlation analysis, for detecting reflections.

Included as appendixes are descriptions of four programs written by the author for the Whirlwind I Digital Computer. These permit high speed computation of: two dimensional polynomial residuals; linear operator prediction errors and their running averages; least squares linear operator coefficients (by an iterative method); and auto-correlations, cross-correlations and "traveling" auto- or cross-correlations.

Thesis Supervisor: Patrick M. Hurley
Title: Professor of Geology

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The author wishes to express his gratitude to Professor P.M. Hurley, thesis supervisor, for his encouragement of and confidence in the work of this thesis.

Among his many co-workers, the author must acknowledge the help of the following persons: Enders A. Robinson for his stimulation, advice, and training in the fundamentals and practice of time series analysis, and in the programming for digital computers; Theodore R. Madden for his interest and assistance, especially in developing the concept of velocity separation and suggesting the problem of least squares residual anomalies; Markwick Smith for many helpful discussions and comments on the various subjects treated here; Howard Briscoe for assistance with programming difficulties; Whirlwind tape room personnel Margaret Mackey and Katherine Campbell, and machine operator Robert Gilday, for their invaluable cooperation.

Finally, the author is deeply indebted to his wife, Pauli, on whom fell the task of preparing the manuscript with its very involved mathematical notations.

INTRODUCTION

It is well known that experimental data taken in Geophysical studies surpasses in accuracy the interpretation that must be made on the data. The reason is that the problems are very complex. For one thing, it can be shown, in the treatment of certain types of problems, that no unique solution exists. An example is the infinity of possible mass distribution corresponding to a given gravity profile. In other problems the physical situation dealt with is so inhomogeneous and anisotropic that exact solution is impossible. It would be hopeless to attempt to explain rigorously the presence of any particular oscillation on a seismogram.

Data such as this, subject to a certain amount of randomness, and on which "best" estimates must be made, is well suited to statistical evaluation. The numerical data taken in gravity surveys does undergo evaluation of this type. The least squares approach, however, is not being utilized on a large scale. This is probably due to practical limitations, and it is one of the problems of this paper to see if these limitations may be minimized.

On the other hand, the raw data of seismology occurs in analogue or curve form. Standard procedures of interpretation consist mainly of rules of thumb, learned by long experience, and still largely dependent on the qualification of the individual interpreter. There is a need to put these procedures on a more rigorous basis. This basis may be found in the concepts of time series as developed in economics, meteorology, and other fields. Much work must be done to determine the best means of applying these concepts to seismic data, since, in certain ways, both the data and the desired information are unique. Another purpose of this paper, then, is to propose several special methods of application, and to develop certain theory necessary for a better understanding of time series concepts as they apply to seismology.

Statistical methods, in general, require computation, and often on a large scale. A "program" written for a digital computer is a tool which will do this work automatically. The author has written several programs for the Whirlwind I Digital Computer to perform computations related to the above discussed problems, and includes these programs as appendixes, with the feeling that other investigators may find them useful.

PART I

LEAST SQUARES RESIDUAL GRAVITY

Introduction

Variations in the attraction of gravity over the surface of the earth are due to many causes, but these often fall into two general categories. Phenomena such as the thickening or thinning of the crust cause relatively slow, smooth and widespread gravity fluctuations. We call these regional effects. On the other hand, such things as ore body emplacements, caverns, and local density heterogeneities cause more rapid irregular changes, and these are termed local effects.

The actual gravity values measured over an area usually represent a combination of regional and local effects. The separation of these effects is of primary importance in interpretation, and many mathematical methods have been devised to eliminate guesswork in the problem. Essentially, most of the methods represent an averaging process which gives at each point and approximate value of the regional effect alone. The local effect is then found simply by subtraction from the measured values.

Many of these methods possess two undesirable aspects. First of all the averaging must be done at each point individually. Secondly, the averaging includes only gravity values in the vicinity of the point considered. It is hard to say just how serious these drawbacks are, but it seems worth-while to investigate a method which does not encounter them. In a least squares approximation all values are averaged simultaneously. Moreover, the resulting approximation is not merely a set of discreet points but a continuous surface of values over the area, a property which is sometimes of value.

The purpose of Part I is to consider in some detail how the method of least squares may be applied to this problem, and how a simplified method of procedure may be set up for practical application.

Part I represents an extension of the work done by W.B. Agocs ‡ . Agocs approximates the regional anomaly by a plane surface derived from least squares criteria. He shows that, in an artifical example, the residual anomaly is better derived from least squares procedures than by the use of the "arithmetic mean regional" procedure. For higher order polynomials than a plane surface the algebra rapidly becomes more involved.

Theory

It is easiest to illustrate the method for an idealized geologic example in two dimensions. Fig. 1.1 shows a wave in the bottom of the crust and a single ore body emplacement. The points on the graph would then be the measured values of gravity across the area. Fitting a fairly low order polynomial to these values by least squares gives us the curve AB which best fits all the points.

‡ Ref. 1

This curve will approximate the regional effect more closely than the local effect, and it is apparent that the fit will be closest at some distance from the ore body. Thus the dashed line of Fig. 1.1, representing the difference between the polynomial and the observed values, gives a good indication of the location of the anomalous mass.

In the two-dimensional problem the approximating polynomial is a surface, and interpretation is made from contours of the residuals.

Let us approximate the regional gravity by a polynomial of order n in x and y.

$$G(xy) = \sum_{i=0}^{n} \sum_{j=0}^{n-i} c_{ij} x^i y^j \qquad 1.1$$

Thus for n=2

$$G(xy) = c_{00} + c_{10}x + c_{20}x^2 + c_{11}xy + c_{01}y + c_{02}y^2$$

The c's are unknown coefficients to be determined in accordance with the condition that the sum of the squares of the residuals is to be minimized. Let g(xy) be the measured values of gravity. Then the residuals are

$$R(xy) = g(xy) - G(xy)$$

and

$$R^2(xy) = [g(xy) - G(xy)]^2$$

Hence

$$\sum_{xy} R^2(xy) = \sum_{xy} [\sum_{i=0}^{n} \sum_{j=0}^{n-i} (c_{ij} x^i y^j \sum_{k=0}^{n} \sum_{\ell=0}^{n-k} c_{k\ell} x^k y^\ell)]$$

$$-2 \sum_{xy} g(xy) \sum_{i=0}^{n} \sum_{j=0}^{n-i} c_{ij} x^i y^j + \sum_{xy} g^2(xy)$$

or

$$\sum_{xy} R^2(xy) = \sum_{xy} \Big[\sum_{i=0}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{n} \sum_{\ell=0}^{n-k} c_{ij} c_{k\ell} x^{k+i} y^{\ell+j} \Big]$$

$$-2 \sum_{xy} \Big[g(xy) \sum_{i=0}^{n} \sum_{j=0}^{n-i} c_{ij} x^i y^j \Big] + \sum_{xy} g^2(xy)$$

Differentiating this expression with respect to c_{ij}, and setting each derivative equal to zero for minimization, we obtain $(n + 1)(n + 2)/2$ linear equations for the same number of unknown coefficients

$$\sum_{k=0}^{n} \sum_{\ell=0}^{n-k} c_{k\ell} \sum_{xy} x^{k+i} y^{\ell+j} = \sum_{xy} g(xy) x^i y^j \qquad 1.2$$

where $j = 0, 1, \ldots\ldots (n-i)$

$i = 0, 1, \ldots\ldots\ldots n$

There are really three variables, or sets of variables, in equations 1.2 — the order of the polynomial n, the set of points xy, and the set of gravity values at these points. The first two of these variables determine the coefficient matrix of the $c_{k\ell}$'s. Once these two are chosen, a unique inverse matrix exists, which, if found, may be used to compute the $c_{k\ell}$'s for all sets of gravity values taken over the same xy pattern. This alone would be a major simplification if the method were to be used on a production basis. But we shall also see that, using a simple reasonable restriction, both the problem of finding the inverse and the form of the inverse itself will be greatly simplified.

Fig. 1.1

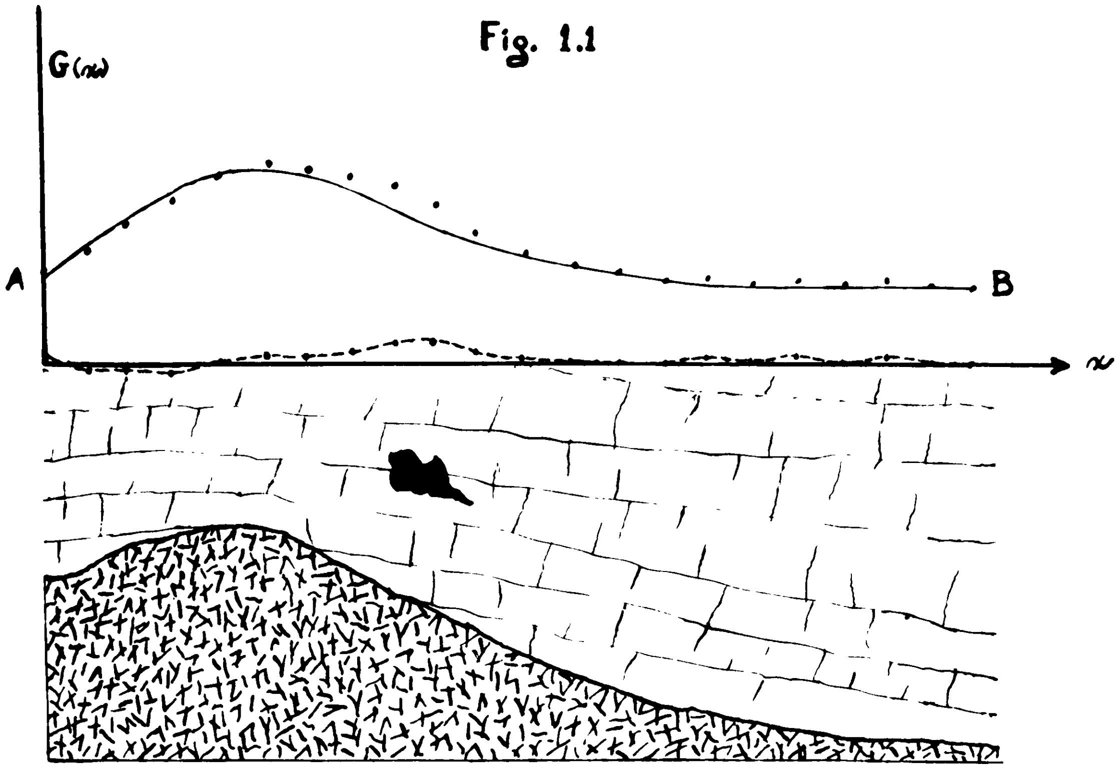
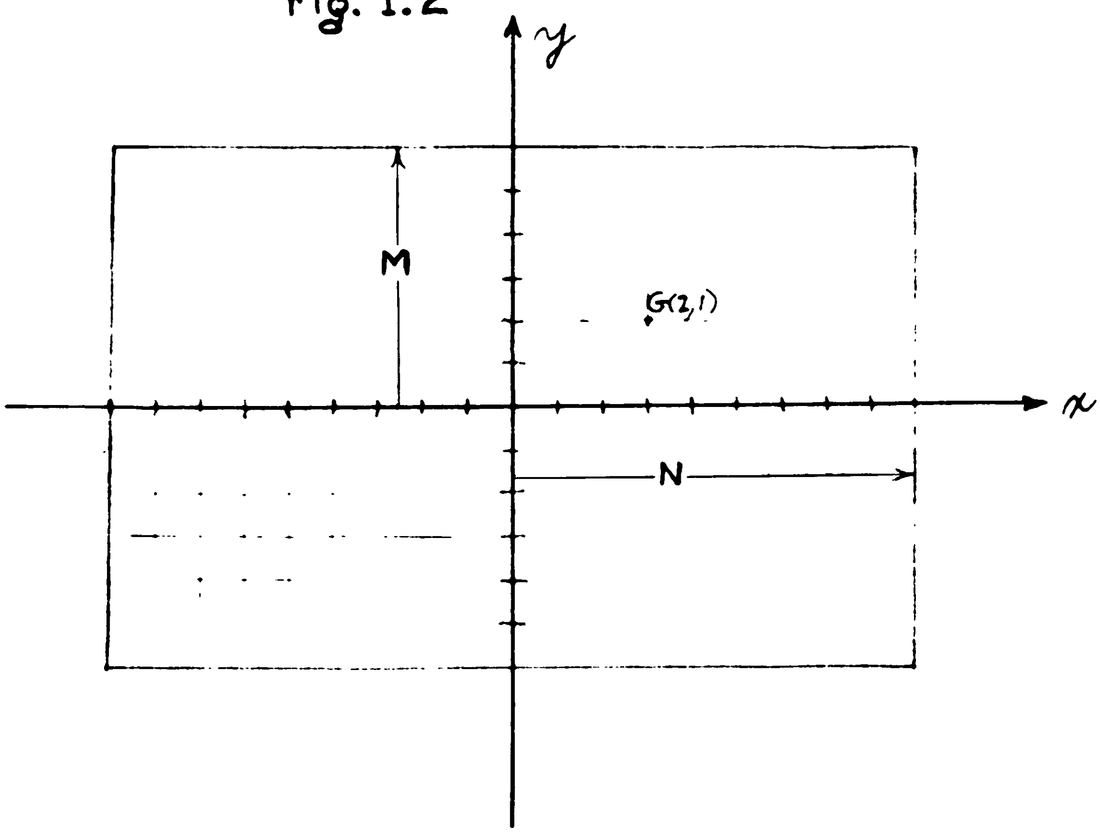


Fig. 1.2



Simplified Solutions

In many cases gravity readings are taken over a square, or at least rectangular, network. When this is so we may take the axes so that the rectangle is symmetrical about them, and number our ordinates and absiscae in integers, as in Fig. 1.2 . It is then easy to see that over such a network summations of the form $x^i y^j$ will vanish whenever i or j is odd. Thus many of the coefficients of the $c_{k\ell}$'s in equations 1.2 will drop out. This leads to considerable simplification, with the bigger systems breaking up into several smaller ones. Furthermore, if we take a definite network we may solve the equations explicitly for the $c_{k\ell}$'s in terms of the summations $\sum_{xy} g(xy) x^i y^j$.

To demonstrate how this is done we shall solve the equations for n = 1, 2, 3, and 4, over a square network of 121 points. The systems are positive definite and symmetric, well adapted to solution by the matrix method of P.D. Crout‡ .

The non-vanishing summations over this network are

$$\Sigma x^0 y^0 \equiv M = 121$$

$$\Sigma x^2 = \Sigma y^2 = 1210 \qquad \Sigma x^2 y^2 = 12100$$

$$\Sigma x^4 = \Sigma y^4 = 21538 \qquad \Sigma x^2 y^4 = \Sigma x^4 y^2 = 215380$$

$$\Sigma x^6 = \Sigma y^6 = 451330 \qquad \Sigma x^2 y^6 = \Sigma x^6 y^2 = 4513300$$

$$\Sigma x^8 = \Sigma y^8 = 10185538 \qquad \Sigma x^4 y^4 = 3833764$$

‡ Ref. 2

Case n = 1

The normal equations are

$$\begin{aligned} c_{00}M &+ c_{10}\Sigma x + c_{01}\Sigma y = \Sigma g(xy) \\ c_{00}\Sigma x &+ c_{10}\Sigma x^2 + c_{01}\Sigma xy = \Sigma g(xy)x \\ c_{00}\Sigma y &+ c_{10}\Sigma xy + c_{01}\Sigma y^2 = \Sigma g(xy)y \end{aligned}$$

Reducing immediately to

$$c_{00} = \frac{\Sigma g(xy)}{M} \qquad c_{10} = \frac{\Sigma g(xy)x}{\Sigma x^2} \qquad c_{01} = \frac{\Sigma g(xy)y}{\Sigma y^2}$$

giving

$$G(xy) = \frac{1}{121}\Sigma g(xy) + \frac{x}{1210}\Sigma g(xy)x + \frac{y}{1210}\Sigma g(xy)y$$

or, to six places

$$G(xy) = 8.26448 \cdot 10^{-4}[10\Sigma g(xy) + x\Sigma g(xy)x + y\Sigma g(xy)y]$$

Case n = 2

The normal equations are

c_{00}	c_{01}	c_{02}	c_{11}	c_{10}	c_{20}		
M	Σy	Σy^2	Σxy	Σx	Σx^2	=	$\Sigma g(xy)$
Σy	Σy^2	Σy^3	Σxy^2	Σxy	Σx^2y	=	$\Sigma g(xy)y$
Σy^2	Σy^3	Σy^4	Σxy^3	Σxy^2	Σx^2y^2	=	$\Sigma g(xy)y^2$
Σxy	Σxy^2	Σxy^3	Σx^2y^2	Σx^2y	Σx^3y	=	$\Sigma g(xy)xy$
Σx	Σxy	Σxy^2	Σx^2y	Σx^2	Σx^3	=	$\Sigma g(xy)x$
Σx^2	Σx^2y	Σx^2y^2	Σx^3y	Σx^3	Σx^4	=	$\Sigma g(xy)x^2$

which reduce to

$$c_{01} = \frac{1}{1210}\Sigma g(xy)y \qquad c_{11} = \frac{1}{12100}\Sigma g(xy)xy \qquad c_{10} = \frac{1}{1210}\Sigma g(xy)x$$

and three equations for c_{00}, c_{02}, and c_{20}

$$121c_{00} + 1210c_{02} + 1210c_{20} = \Sigma g(xy)$$

$$1210c_{00} + 21538c_{02} + 12100c_{20} = \Sigma g(xy)y^2$$

$$1210c_{00} + 12100c_{02} + 21538c_{20} = \Sigma g(xy)x^2$$

The solutions are

$$c_{00} = \frac{1}{9438}[278\Sigma g(xy) - 10(\Sigma g(xy)x^2 + \Sigma g(xy)y^2)]$$

$$c_{02} = \frac{1}{9438}[\Sigma g(xy)y^2 - 10\Sigma g(xy)]$$

$$c_{20} = \frac{1}{9438}[\Sigma g(xy)x^2 - 10\Sigma g(xy)]$$

Thus

$$\begin{aligned} G(xy) = {} & \frac{1}{9438}[278\Sigma g(xy) - 10(\Sigma g(xy)x^2 + \Sigma g(xy)y^2)] \\ & + y[\frac{1}{1210}\Sigma g(xy)y] \\ & + y^2[\frac{1}{9438}(\Sigma g(xy)y^2 - 10\Sigma g(xy))] \\ & + xy[\frac{1}{12100}\Sigma g(xy)xy] \\ & + x[\frac{1}{1210}\Sigma g(xy)x] \\ & + x^2[\frac{1}{9438}\Sigma g(xy)x^2 - 10\Sigma g(xy)] \end{aligned}$$

Or to six places

$$\begin{aligned} G(xy) = {} & [.0294554\Sigma g(xy) - 1.05955\ 10^{-3}(\Sigma g(xy)x^2 + \Sigma g(xy)y^2] \\ & + y[8.26448\ 10^{-4}\Sigma g(xy)y] \\ & + y^2[1.05955\ 10^{-4}(\Sigma g(xy)y^2 - 10\Sigma g(xy))] \\ & + xy[8.26448\ 10^{-5}\Sigma g(xy)xy] \\ & + x[8.26448\ 10^{-4}\Sigma g(xy)x] \\ & + x^2[1.05955\ 10^{-4}(\Sigma g(xy)x^2 - 10\Sigma g(xy))] \end{aligned}$$

Case n = 3

The normal equations are

	c_{00}	c_{01}	c_{02}	c_{03}	c_{11}	c_{12}	c_{21}	c_{10}	c_{20}	c_{30}	
1)	M	y	y^2	y^3	xy	xy^2	x^2y	x	x^2	x^3	$= g$
2)	y	y^2	y^3	y^4	xy^2	xy^3	x^2y^2	xy	x^2y	x^3y	$= gy$
3)	y^2	y^3	y^4	y^5	xy^3	xy^4	x^2y^3	xy^2	x^2y^2	x^3y^2	$= gy^2$
4)	y^3	y^4	y^5	y^6	xy^4	xy^5	x^2y^4	xy^3	x^2y^3	x^3y^3	$= gy^3$
5)	xy	xy^2	xy^3	xy^4	x^2y^2	x^2y^3	x^3y^2	x^2y	x^3y	x^4y	$= gxy$
6)	xy^2	xy^3	xy^4	xy^5	x^2y^3	x^2y^4	x^3y^3	x^2y^2	x^3y^2	x^4y^2	$= gxy^2$
7)	x^2y	x^2y^2	x^2y^3	x^2y^4	x^3y^2	x^3y^3	x^4y^2	x^3y	x^4y	x^5y	$= gx^2y$
8)	x	xy	xy^2	xy^3	x^2y	x^2y^2	x^3y	x^2	x^3	x^4	$= gx$
9)	x^2	x^2y	x^2y^2	x^2y^3	x^3y	x^3y^2	x^4y	x^3	x^4	x^5	$= gx^2$
10)	x^3	x^3y	x^3y^2	x^3y^3	x^4y	x^4y^2	x^5y	x^4	x^5	x^6	$= gx^3$

Summations are assumed for all these quantities and g is written for g(xy). The equations reduce considerably. Equation 5 gives us

$$c_{11} = \frac{1}{12100}\Sigma gxy$$

1, 3, and 9, combine to give three equations for c_{00}, c_{02}, and c_{20}, which have the same solutions as for the case n = 2. 2, 4, and 7, and 6, 8, and 10, combine to give two independent systems which have identical coefficients. Thus 2, 4, 7, are

$$1210c_{01} + 21538c_{03} + 12100c = gy$$
$$21538c_{01} + 451330c_{03} + 215380c_{21} = gy^3$$
$$12100c_{01} + 215380c_{03} + 215380c_{21} = gx^2y$$

With solutions

$$c_{01} = \frac{1}{679536}[4450\Sigma gy - 178\Sigma gy^3 - 72\Sigma gx^2y]$$

$$c_{03} = \frac{1}{679536}[10\Sigma gy^3 - 178\Sigma gy]$$

$$c_{21} = \frac{1}{94380}[\Sigma gx^2y - 10\Sigma gy]$$

also

$$c_{10} = \frac{1}{679536}[4450\Sigma gx - 178\Sigma gx^3 - 72\Sigma gxy^2]$$

$$c_{30} = \frac{1}{679536}[10\Sigma gx^3 - 178\Sigma gx]$$

$$c_{12} = \frac{1}{94380}[\Sigma gxy^2 - 10\Sigma gx]$$

and from the case n = 2

$$c_{00} = \frac{1}{9438}[278\Sigma g - 10(\Sigma gx^2 + \Sigma gy^2)]$$

$$c_{02} = \frac{1}{9438}[\Sigma gy^2 - 10\Sigma g]$$

$$c_{20} = \frac{1}{9438}[\Sigma gx^2 - 10\Sigma g]$$

We also have

$$c_{11} = \frac{1}{12100}\Sigma gxy$$

To six places

$$
\begin{aligned}
G(xy) = \\
& + [.0294554\Sigma g - 1{:}05955\ 10^{-3}(\Sigma gx^2 + \Sigma gy^2)] \\
& + y[6.54859\ 10^{-3}\Sigma gy - 2.61943\ 10^{-4}\Sigma gy^3 - 1.05955\ 10^{-3}\Sigma gx^2 y \\
& + y^2[1.05955\ 10^{-4}(\Sigma gy^2 - 10\Sigma g)] \\
& + y^3[1.47159\ 10^{-5}\Sigma gy^3 - 2.61943\ 10^{-4}\Sigma gy] \\
& + xy[8.26445\ 10^{-5}\Sigma gxy] \\
& + xy^2[1.05955\ 10^{-5}\Sigma gxy^2 - 1.05955\ 10^{-4}\Sigma gx] \\
& + x^2y[1.05955\ 10^{-5}\Sigma gx^2y - 1.05955\ 10^{-4}\Sigma gy] \\
& + x[6.54859\ 10^{-3}\Sigma gx - 2.61943\ 10^{-4}\Sigma gx^3 - 1.05955\ 10^{-3}\Sigma gxy^2] \\
& + x^2[1.05955\ 10^{-4}(\Sigma gx^2 - 10\Sigma g)] \\
& + x^3[1.47159\ 10^{-5}\Sigma gx^3 - 2.61943\ 10^{-4}\Sigma gx]
\end{aligned}
$$

<u>Case n = 4</u>

The normal equations are

	c_{00}	c_{01}	c_{02}	c_{03}	c_{04}	c_{11}	c_{12}	c_{13}	c_{21}	c_{22}	c_{31}	c_{10}	c_{20}	c_{30}	c_{40}		
1)	M	y	y^2	y^3	y^4	xy	xy^2	xy^3	x^2y	x^2y^2	x^3y	x	x^2	x^3	x^4	=	g
2)	y	y^2	y^3	y^4	y^5	xy^2	xy^3	xy^4	x^2y^2	x^2y^3	x^3y^2	xy	x^2y	x^3y	x^4y	=	gy
3)	y^2	y^3	y^4	y^5	y^6	xy^3	xy^4	xy^5	x^2y^3	x^2y^4	x^3y^3	xy^2	x^2y^2	x^3y^2	x^4y^2	=	gy^2
4)	y^3	y^4	y^5	y^6	y^7	xy^4	xy^5	xy^6	x^2y^4	x^2y^5	x^3y^4	xy^3	x^2y^3	x^3y^3	x^4y^3	=	gy^3
5)	y^4	y^5	y^6	y^7	y^8	xy^5	xy^6	xy^7	x^2y^5	x^2y^6	x^3y^5	xy^4	x^2y^4	x^3y^4	x^4y^4	=	gy^4
6)	xy	xy^2	xy^3	xy^4	xy^5	x^2y^2	x^2y^3	x^2y^4	x^3y^2	x^3y^3	x^4y^2	x^2y	x^3y	x^4y	x^5y	=	gxy
7)	xy^2	xy^3	xy^4	xy^5	xy^6	x^2y^3	x^2y^4	x^2y^5	x^3y^3	x^3y^4	x^4y^3	x^2y^2	x^3y^2	x^4y^2	x^5y^2	=	gxy^2
8)	xy^3	xy^4	xy^5	xy^6	xy^7	x^2y^4	x^2y^5	x^2y^6	x^3y^4	x^3y^5	x^4y^4	x^2y^3	x^3y^3	x^4y^3	x^5y^3	=	gxy^3
9)	x^2y	x^2y^2	x^2y^3	x^2y^4	x^2y^5	x^3y^2	x^3y^3	x^3y^4	x^4y^2	x^4y^3	x^5y^2	x^3y	x^4y	x^5y	x^6y	=	gx^2y
10)	x^2y^2	x^2y^3	x^2y^4	x^2y^5	x^2y^6	x^3y^3	x^3y^4	x^3y^5	x^4y^3	x^4y^4	x^5y^3	x^3y^2	x^4y^2	x^5y^2	x^6y^2	=	gx^2y^2
11)	x^3y	x^3y^2	x^3y^3	x^3y^4	x^3y^5	x^4y^2	x^4y^3	x^4y^4	x^5y^2	x^5y^3	x^6y^2	x^4y	x^5y	x^6y	x^7y	=	gx^3y
12)	x	xy	xy^2	xy^3	xy^4	x^2y	x^2y^2	x^2y^3	x^3y	x^3y^2	x^4y	x^2	x^3	x^4	x^5	=	gx
13)	x^2	x^2y	x^2y^2	x^2y^3	x^2y^4	x^3y	x^3y^2	x^3y^3	x^4y	x^4y^2	x^5y	x^3	x^4	x^5	x^6	=	gx^2
14)	x^3	x^3y	x^3y^2	x^3y^3	x^3y^4	x^4y	x^4y^2	x^4y^3	x^5y	x^5y^2	x^6y	x^4	x^5	x^6	x^7	=	gx^3
15)	x^4	x^4y	x^4y^2	x^4y^3	x^4y^4	x^5y	x^5y^2	x^5y^3	x^6y	x^6y^2	x^7y	x^5	x^6	x^7	x^8	=	gx^4

Equations 1, 3, 5, 10, 13, 15, reduce to give a system of six equations for c_{00}, c_{02}, c_{04}, c_{22}, c_{20}, and c_{40}. 2, 4, 9, and 7, 12, 14, give two sets of equations for c_{01}, c_{03}, c_{21}, and c_{10}, c_{30}, c_{12}, respectively, which are equivalent to the corresponding equations for the case n = 3.

The new equations to be solved are

c_{00}	c_{02}	c_{04}	c_{22}	c_{20}	c_{40}	
M	y^2	y^4	x^2y^2	x^2	x^4	$= \Sigma g$
y^2	y^4	y^6	x^2y^4	x^2y^2	x^4y^2	$= \Sigma gy^2$
y^4	y^6	y^8	x^2y^6	x^2y^4	x^4y^4	$= \Sigma gy^4$
x^2y^2	x^2y^4	x^2y^6	x^4y^4	x^4y^2	x^6y^2	$= \Sigma gx^2y^2$
x^2	x^2y^2	x^2y^4	x^4y^2	x^4	x^6	$= \Sigma gx^2$
x^4	x^4y^2	x^4y^4	x^6y^2	x^6	x^8	$= \Sigma gx^4$

and

c_{11}	c_{13}	c_{31}	
x^2y^2	x^2y^4	x^4y^2	$= \Sigma gxy$
x^2y^4	x^2y^6	x^4y^4	$= \Sigma gxy^3$
x^4y^2	x^4y^4	x^6y^2	$= \Sigma gx^3y$

The last set has solutions

$$c_{11} = \frac{1}{679536} [689.84\Sigma gxy - 17.8(\Sigma gxy^3 + \Sigma gx^3y)]$$

$$c_{13} = \frac{1}{679536} [\Sigma gxy^3 - 17.8\Sigma gxy]$$

$$c_{31} = \frac{1}{679536} [\Sigma gx^3y - 17.8\Sigma gxy]$$

The solution of the first set to six places is

$$c_{00} = 4.53280\ 10^{-2}\Sigma g + 1.58932\ 10^{-4}(\Sigma gy^4 + \Sigma gx^4)$$

$$1.35839\ 10^{-4}\Sigma gx^2y^2 - 6.399122\ 10^{-3}(\Sigma gy^2 + \Sigma gx^2)$$

$$c_{02} = 1.62141\ 10^{-3}\Sigma gy^2 - 2.81527\ 10^{-3}\Sigma g$$

$$-5.51847\ 10^{-5}\Sigma gy^4 - 1.35839\ 10^{-5}(\Sigma gx^2y^2 - 10\Sigma gx^2)$$

$$c_{04} = 2.20739\ 10^{-6}\Sigma gy^4 - 5.51847\ 10^{-5}\Sigma gy^2 + 1.58932\ 10^{-5}\Sigma g$$

$$c_{22} = 1.35839\ 10^{-6}(\Sigma gx^2y^2 - 10(\Sigma gx^2 + \Sigma gy^2) + 100\Sigma g)$$

$$c_{20} = 1.62141\ 10^{-3}\Sigma gx^2 - 2.81527\ 10^{-3}\Sigma g$$

$$-5.51847\ 10\text{-}^{5}\Sigma gx^2 - 1.35839\ 10^{-5}(\Sigma gx^2y^2 - 10\Sigma gy^2)$$

$$c_{40} = 2.20739\ 10^{-6}\Sigma gx^4 - 5.51847\ 10^{-5}\Sigma gx^2 + 1.58932\ 10^{-5}\Sigma g$$

To simplify writing G(xy) we introduce the abbreviations

$A = \Sigma g$	$H = \Sigma gx^3y$
$B = \Sigma gx$	$I = \Sigma gxy^2$
$C = \Sigma gx^2$	$J = \Sigma gx^2y^2$
$D = \Sigma gx^3$	$K = \Sigma gxy^3$
$E = \Sigma gx^4$	$L = \Sigma gy$
$F = \Sigma gxy$	$M = \Sigma gy^2$
$G = \Sigma gx^2y$	$N = \Sigma gy^3$
	$P = \Sigma gy^4$

Hence

$$
\begin{aligned}
G(xy) = \; & [4.53280\ 10^{-2}A + 1.58932\ 10^{-4}(P + E) \\
& \quad +1.35839\ 10^{-4}J - 6.39122\ 10^{-3}(M + C)] \\
& + y[6.54859\ 10^{-3}L - 2.61943\ 10^{-4}N - 1.05955\ 10^{-3}G] \\
& + y^2[1.62141\ 10^{-3}M - 2.81527\ 10^{-3}A - 5.51847\ 10^{-5}P \\
& \quad -1.35839\ 10^{-5}(J - 10C)] \\
& + y^3[1.47159\ 10^{-5}N - 2.61943\ 10^{-4}L] \\
& + y^4[2.20739\ 10^{-6}P - 5.51847\ 10^{-5}M + 1.58932\ 10^{-5}A] \\
& + xy[1.01516\ 10^{-3}F - 2.61943\ 10^{-5}(K + H)] \\
& + xy^2[1.05955\ 10^{-5}I - 1.05955\ 10^{-4}B] \\
& + xy^3[1.47159\ 10^{-6}K - 2.61943\ 10^{-5}F] \\
& + x^2y[1.05955\ 10^{-5}G - 1.05955\ 10^{-4}B] \\
& + x^2y^2[1.35839\ 10^{-6}(J - 10(C + M) + 100A)] \\
& + x^3y[1.47159\ 10^{-6}H - 2.61943\ 10^{-5}F] \\
& + x[6.54859\ 10^{-3}B - 2.61943\ 10^{-4}D - 1.05955\ 10^{-3}I] \\
& + x^2[1.62141\ 10^{-3}C - 2.81527\ 10^{-3}A - 5.51847\ 10^{-5}C \\
& \quad -1.35839\ 10^{-5}(J - 10M)] \\
& + x^3[1.47159\ 10^{-5}D - 2.61943\ 10^{-4}B] \\
& + x^4[2.20739\ 10^{-6}E - 5.51847\ 10^{-5}C - 1.58932\ 10^{-5}A]
\end{aligned}
$$

Discussion

An interesting property which has developed in these four cases makes the extension to higher order approximations somewhat simpler. If n is odd then all the coefficients c_{ij} whose subscripts add to an even number are the same as the corresponding coefficients for the (n - 1)rst case. If n is even the coefficients with subscripts adding to an odd number are the same as for the preceding case. This property can be shown directly from equation 1.2.

Thus for the case n = 5 we expect nine of the coefficients (c_{00}, c_{02}, c_{04}, c_{40}, c_{20}, c_{11}, c_{13}, c_{31}) to be the same as for the case n = 4, and we need only write the twelve remaining equations for i + j odd.

A polynomial of order equal to the number of points taken will exactly fit the data. However it is practically impossible to use polynomials even approaching such a high order for reasonably-sized gridworks, and this danger seems slight. There is still a real problem in the choice of n. If the regional effect is in reality a fairly low order effect, polynomials of high n will begin to approximate the local anomalies too closely. Other systems, however, run into the same problem, and this point would be best settled by experience with the data.

Another important practical consideration is the amount of work to be done, i.e., the determination of the summations $\Sigma gx^i y^j$, of the c_{ij}'s and the solution of $G(xy)$ at each point. We devote the next section to this problem.

Applications

To illustrate the work necessary we discuss a convenient scheme for application to the case $n = 3$. The use of a computing machine with cumulative multiplication is desirable.

Assume that the grid has been determined and the gravity values written at each intersection as shown. This is done on tracing paper as shown in Fig. 1.3.

The numbers α_i and β_i above each vertical line and to the left of each horizontal line represent the sums of $g(xy)$ along those lines. Then as we may easily compute the sums Σg, Σgx, Σgx^2, Σgx^3, Σgy, Σgy^2, Σgy^3, from the relations

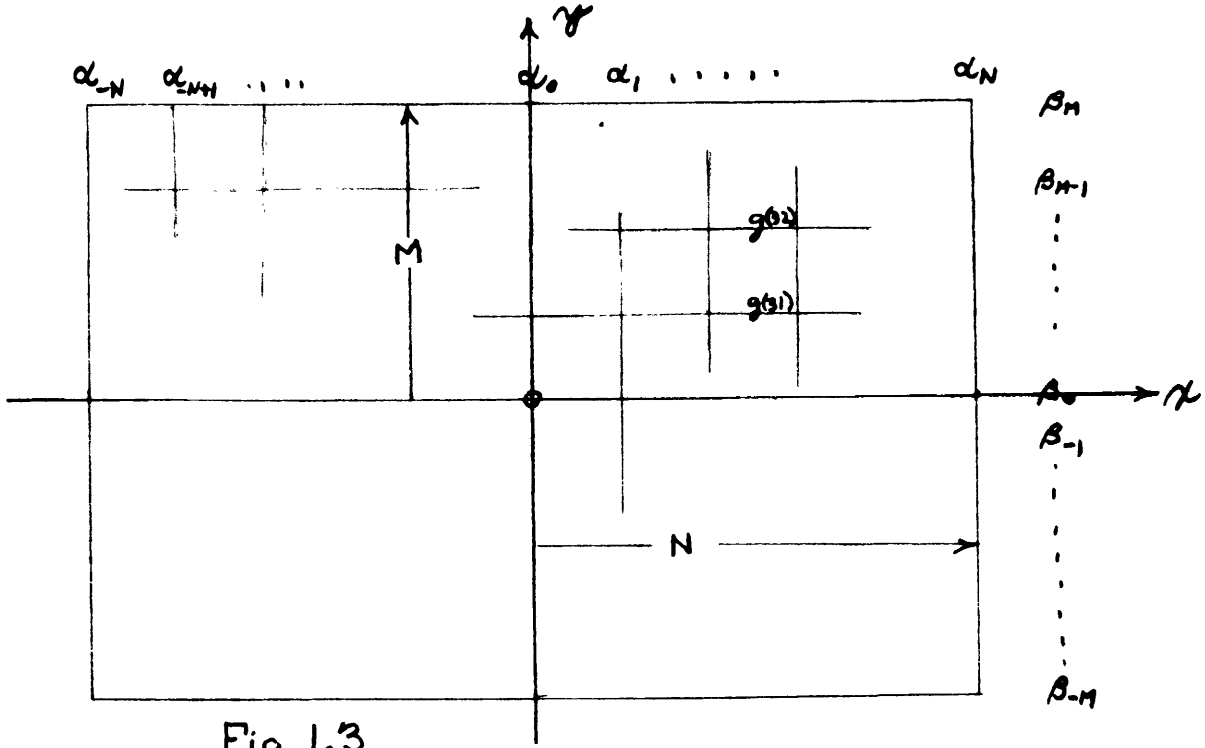
$$\Sigma g = \sum_i \alpha_i \qquad \Sigma gx^3 = \sum_i \alpha_i (i)^3 \qquad gy^2 = \sum_i \beta_i (i)^2$$

$$\Sigma gx = \sum_i \alpha_i (i) \qquad \Sigma gy = \sum_i \beta_i (i) \qquad gy^3 = \sum_i \beta_i (i)^3$$

$$\Sigma gx^2 = \sum_i \alpha_i (i)^2$$

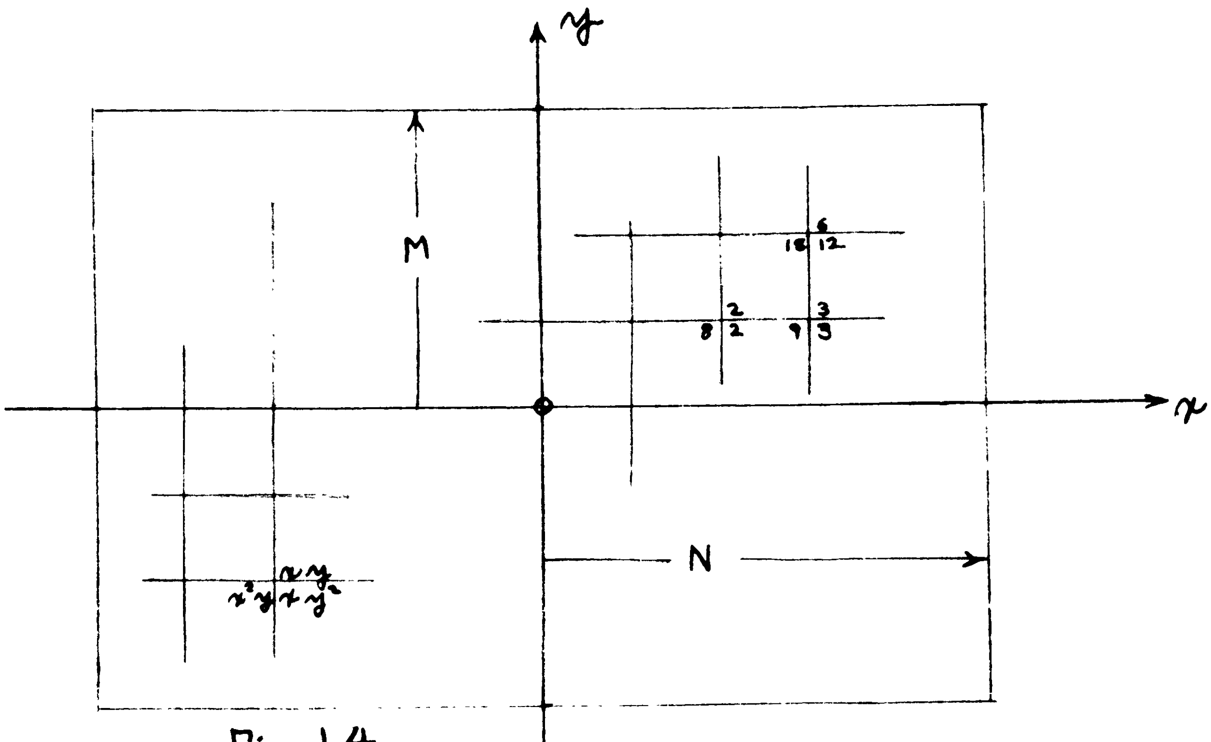
Each of these computations involves one machine operation of eleven cumulative multiplications.

For the remaining summations Σgxy, Σgx^2y, Σgxy^2 it is convenient to have a similar grid which can be placed under the original one. This second grid has the values of xy, x^2y, xy^2, at each point as shown in Fig. 1.4 and can be used for each application.



y
α_{-N} α_{-N+1} ....
α_0 α_1 .....
α_N
β_M
β_{M-1}
M
g(3,2)
g(3,1)
β_0
x
β_{-1}
N
β_{-M}

Fig. 1.3



y
M
6
18 12
2
8 2
3
9 3
x
N
xy
$x^2y + y^2$

Fig. 1.4

The values of g(xy) then appear in the vacant upper left hand corner of each point, making the multiplications apparent. Each of these three summations then involves a cumulative addition of 100 multiplications.

The c_{ij}'s are then found as ten cumulative additions of two of three multiplications each.

G(xy) is now completely determined with the writing down of less than 50 numbers and it remains to solve this equation for each point. This involves ten cumulative multiplications at each of the 121 points with a final subtraction to determine the residuals. A second tracing paper grid laid over both of the others would simplify this and the residuals could be written down in a form ready to be contoured.

A nice feature of this scheme is the absence of any tabulation of data. It may be extended fairly simply to higher degrees.

<u>Example</u>

As a test of this method residuals were computed on gravity readings supplied by a mining company. This data was not in a convenient form for use since the readings were taken in mine tunnels and not over a grid. To get them in grid form, the readings were first contoured as shown in Fig. 1.5 and then values extrapolated to the grid. This involves several inaccuracies. First of all, where readings are sparse, the extrapolated values are bound to contain considerable error. Secondly, the computations weight all values equally so that the accurately extrapolated points, in areas of dense readings, suffer from the inaccuracies in the less dense areas.

Three sets of residuals were computed, one for second, third, and fourth order polynomials. This was done to test the effect of polynomial degree on the residuals. The polynomials were fitted directly to the raw gravity, without making the usual topographic corrections. The justification for neglecting to do this is seen in Fig. 1.11. This figure shows values of regional minus topographic corrections computed by the mining company, and contours of these values. The contours demonstrate that this correction is a low order effect (in this particular situation) and can obviously be easily absorbed into a polynomial as low as degree two.

Figs. 1.6, 1.7, and 1.8 then show residuals for second, third, and fourth order polynomials respectively, and were computed as described previously. Once the reading had been contoured and extrapolated, it took about a day to compute each set of residuals. The computed polynomials appear in the upper right-hand corner of Figs. 1.6, 1.7, and 1.8. Fig. 1.10 shows the contours of the polynomials themselves, and shows that the similarity of the residuals is due to the similarity of the polynomials used.

In Fig. 1.9 is contoured the residual gravity as computed by the mining company which supplied the data. Their computational procedure required several months to produce this diagram, which, in important respects, is quite similar to the contours of Figs. 1.6, 1.7, and 1.8. Part of the difference is due to the fact that the least squares residuals are forced to oscillate around a mean of zero, so that many negative contours appear. Other differences may well be attributable to the inaccuracies of contouring as mentioned above. It seems clear however, that the similarity is sufficiently great to justify the use of the least squares procedure, at least for a first evaluation of gravity data. This seems particularly true in view of the relative speed with which this procedure may be carried out.

These results were encouraging enough so that a program was written for the WWI Digital Computer to perform the majority of the computations automatically. This program finds the residuals for a polynomial up to the sixth order over an arbitrary grid shape, once the polynomial is known. A description of this program appears in Appendix A.

In the next section, we take up the problem of setting up the normal equations for various sizes and shapes of grids.

RAW GRAVITY - 1300.0 CONTOURED

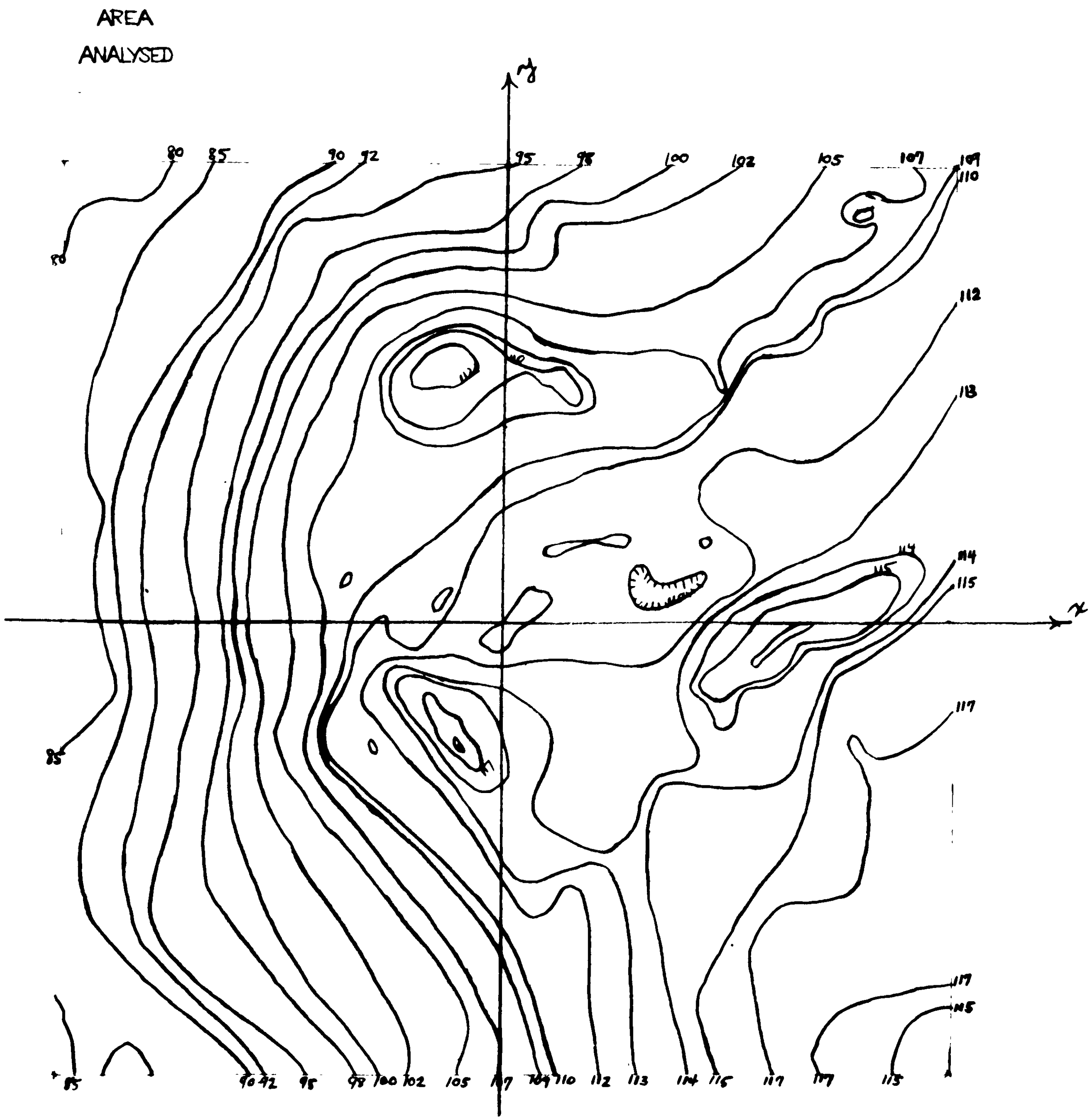


Fig. 1.5

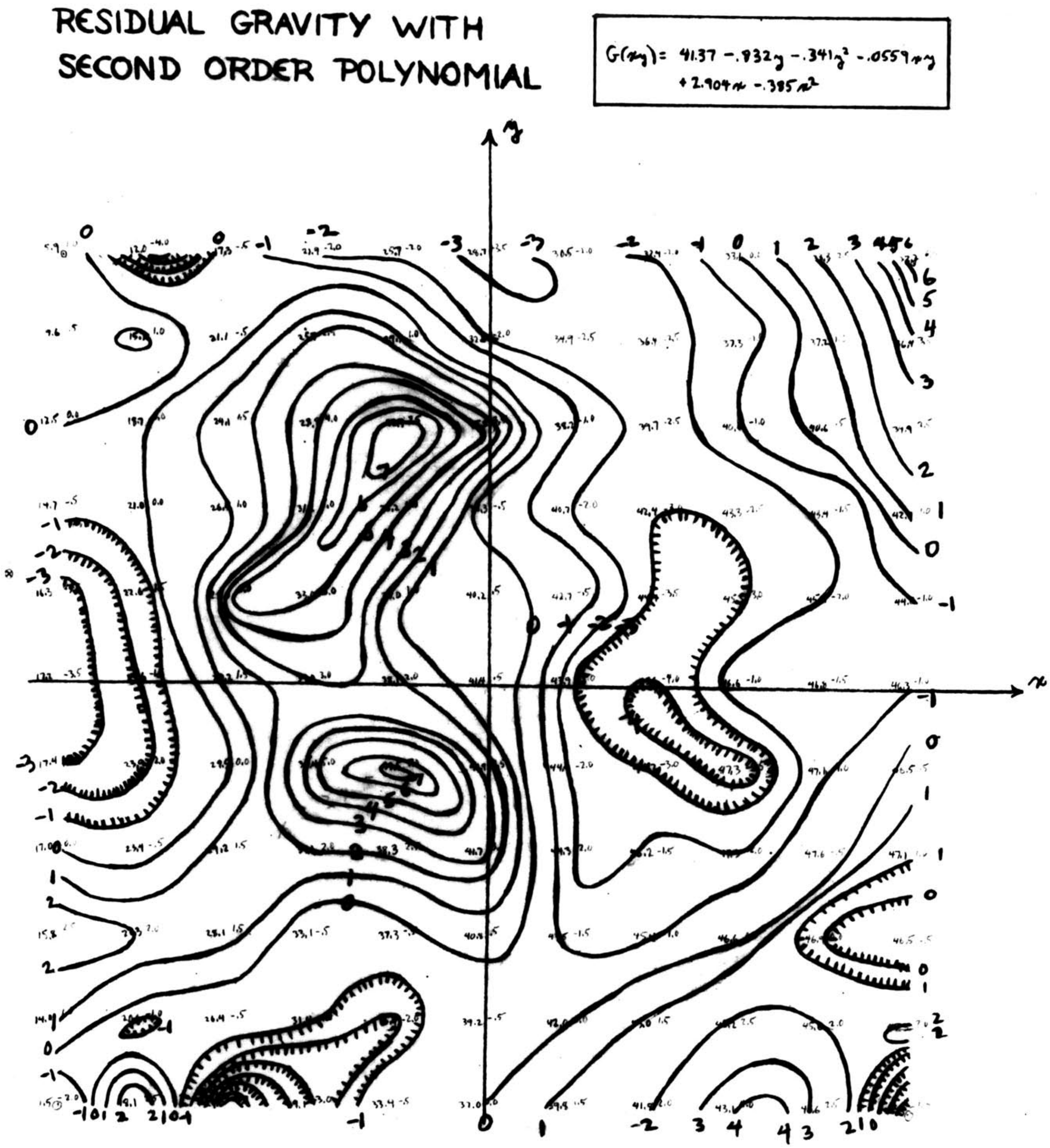

RESIDUAL GRAVITY WITH
SECOND ORDER POLYNOMIAL
G(xy) = 41.37 − .832y − .341y² − .0559xy
+ 2.904x − .385x²

Fig. 1.6

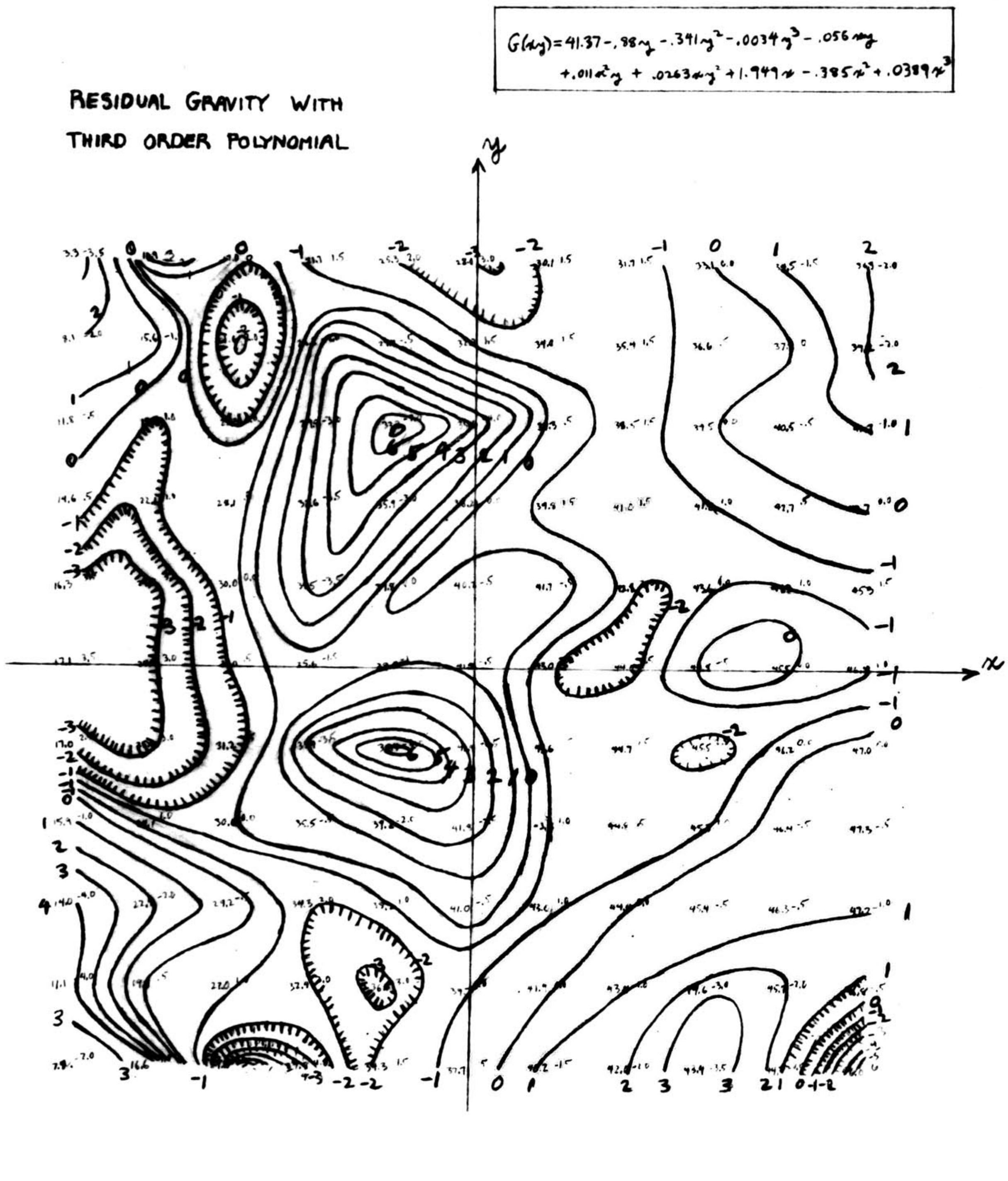

G(x,y)=41.37−.88y−.341y²−.0034y³−.056xy
+.011x²y + .0263xy² + 1.949x − .385x² + .0389x³
RESIDUAL GRAVITY WITH
THIRD ORDER POLYNOMIAL
y
x

Fig. 1.7

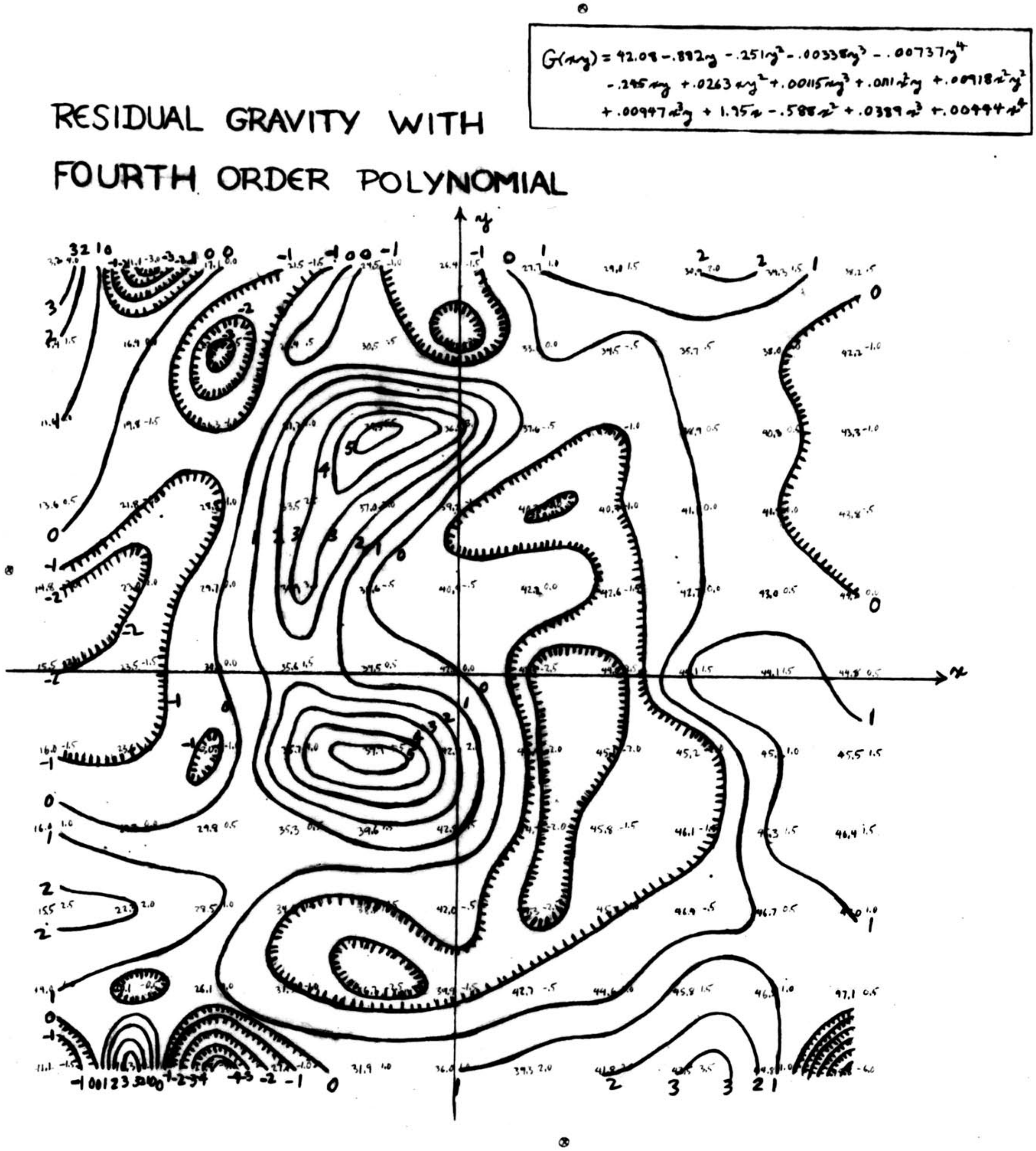

G(xy) = 42.08 - .882y - .251y² - .00338y³ - .00737y⁴
- .245xy + .0263xy² + .00115xy³ + .011x²y + .00918x²y²
+ .00947x³y + 1.95x - .588x² + .0389x³ + .00494x⁴
RESIDUAL GRAVITY WITH
FOURTH ORDER POLYNOMIAL
y
x

Fig. 1.8

RESIDUAL GRAVITY
COMPANY DATA

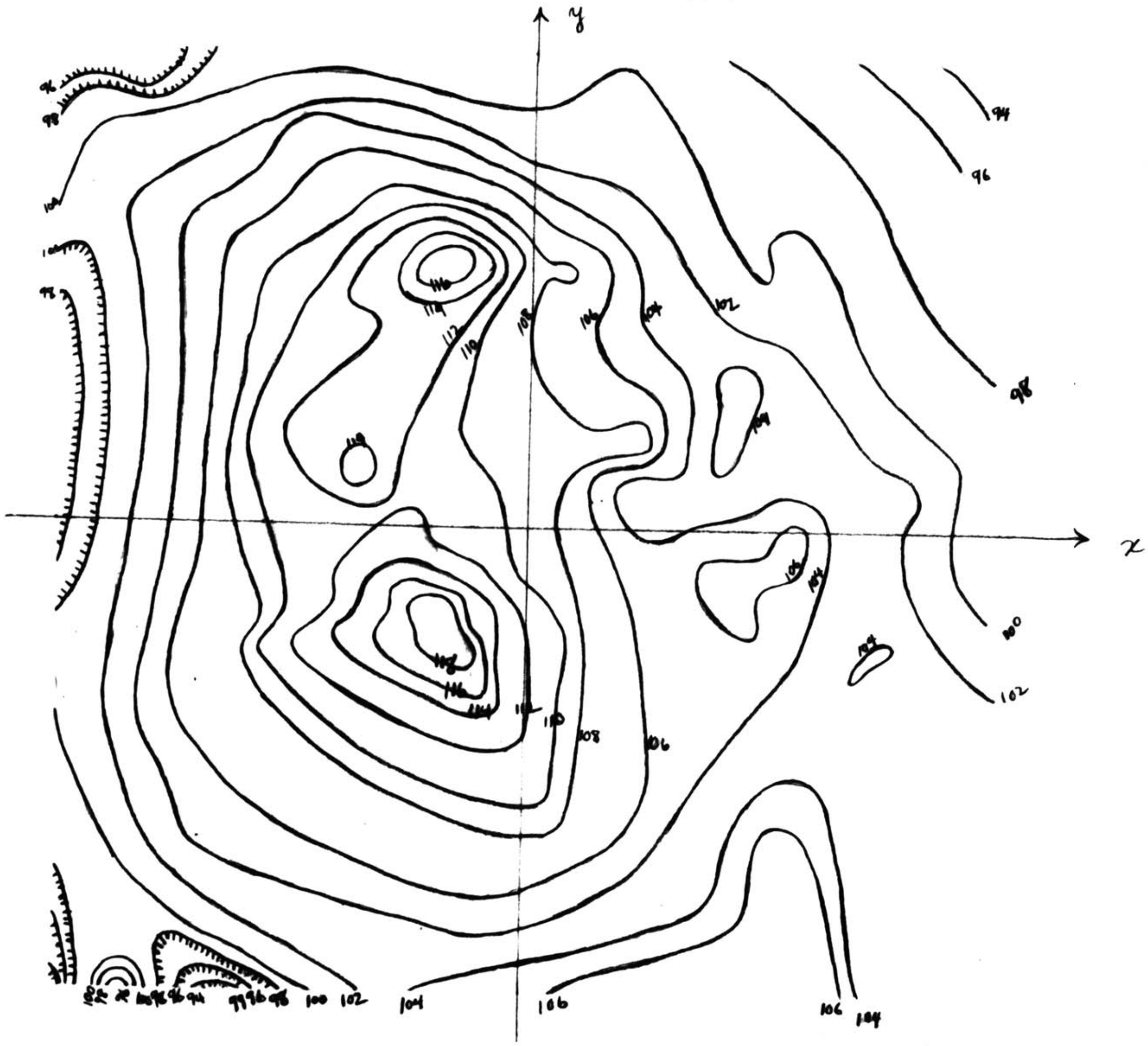
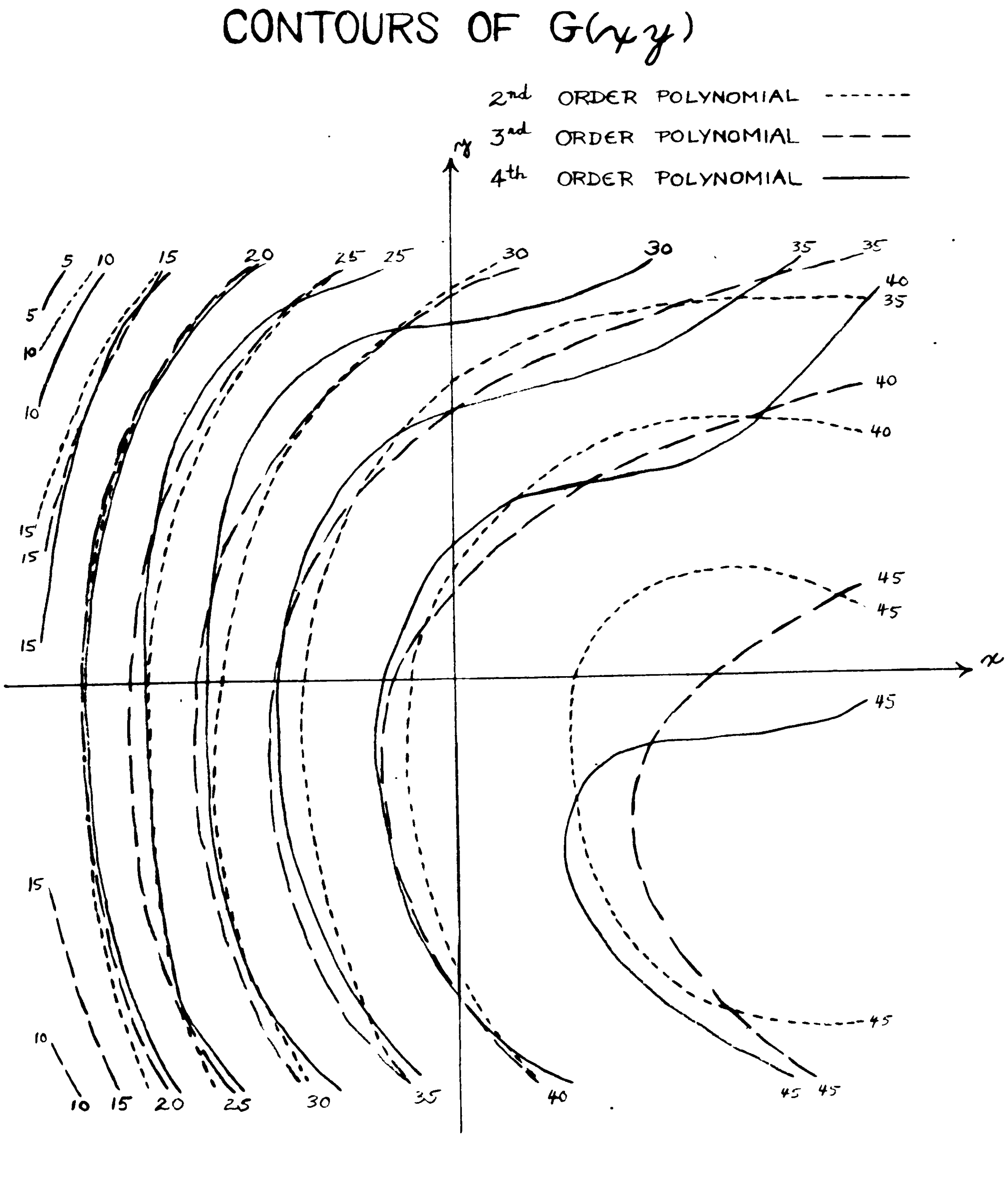


Fig. 1.9



CONTOURS OF G(x,y)
2nd ORDER POLYNOMIAL
3rd ORDER POLYNOMIAL
4th ORDER POLYNOMIAL
y
x

Fig. 1.10

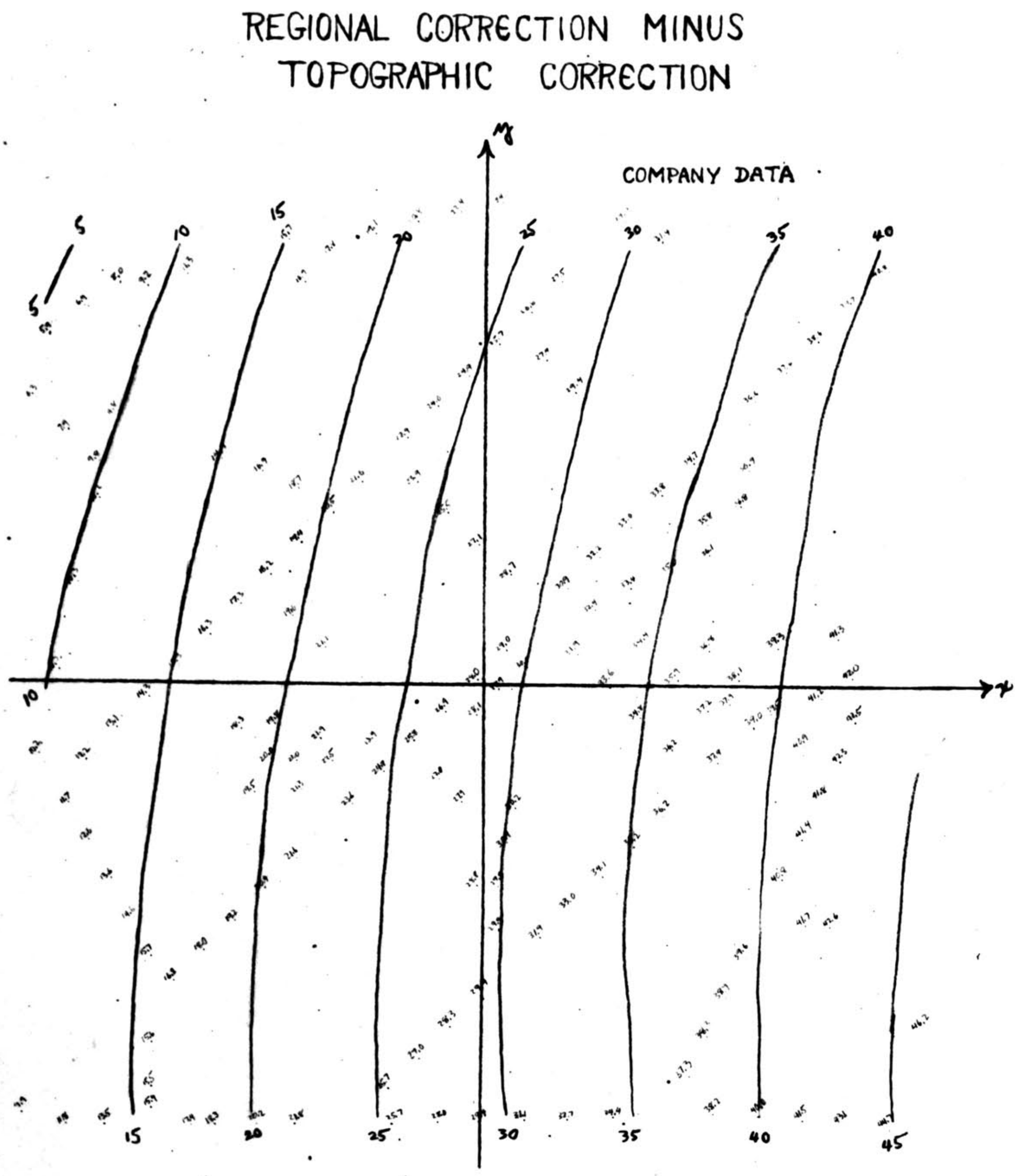

REGIONAL CORRECTION MINUS
TOPOGRAPHIC CORRECTION
COMPANY DATA
y
x

Fig. 1.11

Setting Up the Normal Equations

We are concerned here with the problem of setting up the normal equations for various grids and polynomial degrees. If we limit ourselves to polynomials of degree 4 or less, we are then interested in finding the following quantities:

$$\Sigma x^2 \quad \Sigma x^4 \quad \Sigma x^6 \quad \Sigma x^8 \quad \Sigma y^4 \quad \Sigma y^6 \quad \Sigma y^8$$

$$\Sigma x^2 y^2 \quad \Sigma x^2 y^4 \quad \Sigma x^2 y^6 \quad \Sigma x^4 y^2 \quad \Sigma x^4 y^4 \quad \Sigma x^6 y^2 \quad \Sigma x^0 y^0 \qquad 1.3$$

where the summations are to be taken over the particular grid we are dealing with.

If the grid has the dimensions 2N by 2M as shown in Fig. 1.2, we may set up a fairly simple procedure for finding these summations.

First we note that Σx^k over the grid is equal to the Σx^k on a single horizontal line, times the number of lines. Thus

$$\Sigma x^k = (2M + 1)\sum_{i=-N}^{N} i^k$$

but since in our case k is always even

$$\Sigma x^k = 2(2M + 1)\left(\sum_{i=1}^{N} i^k\right) \qquad 1.4$$

Likewise

$$\Sigma y^k = 2(2N + 1)\sum_{i=1}^{M} i^k$$

For the cross terms $\Sigma x^k y^\ell$ we have

$$\sum_{grid} x^k y^\ell = \sum_{j=-M}^{M} \sum_{i=-N}^{N} i^k j^\ell$$

$$= [\sum_{-N}^{N} i^k] [\sum_{-M}^{M} j^\ell]$$

$$\sum_{grid} x^k y^\ell = 4(\sum_{i=1}^{N} i^k)(\sum_{j=1}^{M} j^\ell)$$

Hence the sums 1.3 are easily derivable from equations 1.4 and 1.5 if we tabulate the quantities $\sum_{i=1}^{L} i^k$. Table I gives values of i^k from which the sums $\sum_{i=1}^{L} i^k$ are derived, and Table II tabulates these sums for L up to 25 and k = 2, 4, 6, 8. The latter Table allows us to compute the sums 1.3 for any grids measuring up to 50 by 50. a grid this size would encompass 2601 gravity readings which seems ample for most applications.

Table III contains the sums 1.3 computed for six representative grids 10 by 10, 10 by 20, 20 by 20, 30 by 30, 40 by 40, and 50 by 50.

TABLE I i^k

for $k = 2, 4, 6, 8,\quad i = 1, 2, 3 \ldots 25$

i	i^2	i^4	i^6	i^8
1	1	1	1	1
2	4	16	64	256
3	9	81	729	6561
4	16	256	4096	65536
5	25	625	15625	390625
6	36	1296	46656	1679616
7	49	2401	117649	5764801
8	64	4096	262144	16777216
9	81	6561	531441	43046721
10	100	10000	1000000	100000000
11	121	14641	1771561	214358881
12	144	20736	2985984	429981696
13	169	28561	4826809	815730721
14	196	38416	7529536	1475789056
15	225	50625	11390625	2562890625
16	256	65536	16777216	4294967296
17	289	83521	24137569	6975757441
18	324	104976	34012224	11019960576
19	361	130321	47045881	16983563041
20	400	160000	64000000	25600000000
21	441	194481	85766121	37822859361
22	484	234256	113379904	54875873536
23	529	279841	148035889	78310985281
24	576	331776	191102976	110075314176
25	625	390625	244140625	152587890625

TABLE II $\sum_{i=1}^{L} i^k$

for k = 2, 4, 6, 8   L = 1, 2, 3, ... 25

L	k=2	k=4	k=6	k=8
1	1	1	1	1
2	5	17	65	257
3	14	98	794	6818
4	30	354	4890	72354
5	55	979	20515	462979
6	91	2275	67171	2142595
7	140	4676	184820	7907396
8	204	8772	446964	24684612
9	285	15333	978405	67731333
10	385	25333	1978405	167731333
11	506	39974	3749966	382090214
12	650	60710	6735950	812071910
13	819	89271	11562759	1627802631
14	1015	127687	19092295	3103591687
15	1240	178312	30482920	5666482312
16	1496	243848	47260136	9961449608
17	1785	327369	71397705	16937207049
18	2109	432345	105409929	27957167625
19	2470	562666	152455810	44940730666
20	2870	722666	216455810	70540730666
21	3311	917147	302221931	108363590027
22	3795	1151403	415601835	163239463563
23	4324	1431244	563637724	241550448844
24	4900	1763020	754740700	351625763020
25	5525	2153645	998881325	504213653645

TABLE III $\Sigma x^k y^\ell$

Over various grids measuring 2N x 2M

for k = 0, 2, 4, 6, 8 ℓ = 0, 2 ...(8-k)

	N=5 M=5	N=5 M=10	N=10 M=10	N=15 M=15	N=20 M=20	N=25 M=25
$\Sigma x^0 y^0$ =	121	231	441	961	1681	2601
Σx^2 =	1210	2310	16170	76880	235340	563550
Σx^4 =	21538	41118	1063986	11055344	59258612	219671790
Σx^6 =	451330	861630	83093010	1889941040	17749376420	101885895150
Σx^8 =	10185538	19445118	7044715986	351321903344	5784339914612	51429792671790
Σy^2 =	1210	8470	16170	76880	235340	563550
Σy^4 =	21538	557326	1063986	11055344	59258612	219671790
Σy^6 =	451330	43524910	83093010	1889941040	17749376420	101885895150
Σy^8 =	10185538	3690089326	7044715986	351321903344	5784339914612	51429792671790
$\Sigma x^2 y^2$ =	12100	84700	592900	6150400	32947600	122102500
$\Sigma x^2 y^4$ =	215380	5573260	39012820	884427520	8296205680	47595554500
$\Sigma x^2 y^6$ =	4513300	435249100	3046743700	151195283200	2484912698800	22075277282500
$\Sigma x^4 y^2$ =	215380	1507660	39012820	884427520	8296205680	47595554500
$\Sigma x^6 y^2$ =	4513300	31593100	3046743700	151195283200	2484912698800	22075277282500
$\Sigma x^4 y^4$ =	3833764	99204028	2567043556	127180677376	2088984590224	18552747144100

PART II

SEISMIC RECORD ANALYSIS BY LINEAR OPERATORS

Introduction

In the study of reflection seismic records, taken in the exploration for oil, it is becoming increasingly difficult to pick reflection times by the standard procedures. The reason for this is that, as the simpler geologic areas are being fully exploited, exploration is being forced into the more complicated areas. Seismic records taken in these structurally complex areas contain much in the way of unwanted information and much not-understood information. Energy reflected from the strata of interest is largely masked by this "noise". At least two different approaches to unscrambling these records are being developed at present.

The first of these approaches is largely instrumentational. Its principle is: take more and more information (more traces on each record, etc.), filter it in different ways and mix it up in a variety of combinations to see if a procedure for averaging out the unwanted information can be arrived at. This approach has led oil companies to the use of 24-trace records, each trace representing the responses from up to thirty geophones. The success of these methods is not publicly available, but the oil industry is expressing great interest in the approach described below, so probably they are not completely satisfactory.

The second of these approaches is basicly analytic. Rather than taking more information, we attempt to sharpen up the interpretive procedure on the information we have.

The search for such procedure has been largely carried out at MIT in the Mathematics Department, and subsequently in the Mathematics and Geology Departments. The tools in this analysis have been the statistics of time series.

Statistical Methods

After some experimentation it was found at MIT that the use of the "linear operator" seemed most promising in the determination of reflection times. The exact methods used are described in Ref's. 5 and 6. The linear operator permits a measure of the change in dynamics as we proceed down a seismic record. As these dymanics are amplitude, frequency, and phase relationships, it was hoped that the dynamical change at a reflection could be discriminated even when the changes due to amplitude were small. This was hoped for since the usual interpretive procedures depend heavily on "amplitude reflections". The results were very encouraging and stimulated increased research.

One direction this study has taken is the empirical one. We know the linear operator gives us added information. But, since there is considerable freedom in the choice of the exact mathematical form of the operator we use, we try many different forms and see which ones give us the most information. This is a trial and error procedure and involves an immense amount of computation. For this reason a program was written for the WWI Digital Computer which would compute automatically the measure of dynamic change, for a great variety of forms of linear operators, at very high speed. A copy of this program and a description of its functions is contained in Appendix B. Case studies designed to test the effects of individual parameters of the linear operator are being run with this program, but the results are as yet incomplete.

Along with this empirical approach an attempt is being made to study the linear operator from theoretical grounds. Although the form of the operator which is being tested at present is relatively complicated, it is instructive to consider a simpler form, the so-called "cosine operator". This operator is a mathematical expression which generates a pure cosine wave of given frequency. We can determine quite simply the effects of this type of operator on various time series including those found on seismic records. We hope to gain insight into the physical function of such operators as well as correspondence between them and simple filters.

We shall also consider two other more practical problems connected with the statistical analysis of seismograms by the use of linear operators. One concerns certain iterative methods for approaching the values of the linear operator coefficients for least squares fitting. The other is a related problem, the necessity for accuracy in finding these values.

Single Frequency Cosine Operators

A cosine operator is a prediction mechanism which exactly predicts equally spaced points on a cosine wave. It has the general form ‡

$$\hat{x}_{i+2} = c + ax_{i+1} + bx_i \qquad 2.1$$

where $a = 2 \cos 2\pi hf \equiv 2u$

$b = -1$

h = time between observation

$c = (1-a-b)\bar{x} = 2(1-u)\bar{x}$

$\bar{x}$ = mean of series

f = frequency of cosine wave

$\hat{x}_{i+2}$ = predicted value of x_{i+2}

Suppose we use this operator to predict an arbitrary series x. Then the error of prediction $x_{i+2} - \hat{x}_{i+2}$ will be

$$\begin{aligned} E_{i+2} &= x_{i+2} - [2(1-u)\bar{x} + 2ux_{i+1} - x_i] \qquad 2.2 \\ &= x_{i+2} - 2(1-u)\bar{x} - 2ux_{i+1} + x_i \\ &= (x_{i+2} - \bar{x}) + (x_i - \bar{x}) - 2u(x_{i+1} - \bar{x}) \end{aligned}$$

For simplicity let us deal with a series X_i measured around its equilibrium mean $\bar{x}$, i.e. $X_i = x_i - \bar{x}$ then 2.2 becomes

$$E_{i+2} = X_{i+2} + X_i - 2uX_{i+1} \qquad 2.3$$

Now if we sum the squares of these errors over an interval of the series we get

‡ Ref. 6

$$\sum_i E_{i+2}^2 = \Sigma(X_{i+2}^2 + X_i^2 + 4u^2X_{i+1}^2 + 2X_{i+2}X_i - 4uX_{i+2}X_{i+1} - 4uX_iX_{i+1} \qquad 2.4$$

If the series is stationary and the interval sufficiently great we may write this in terms of the auto-correlations. Let the series be normalized so that $\sum_i X_i^2 = 1$ then

$$\sum_i E_{i+2}^2 = R_0 + R_0 + 4u^2 R_0 + 2R_2 - 8uR_1 \qquad 2.5$$

where R_i = $i^{\underline{th}}$ lag auto-correlation and $R_0 = 1$

or

$$\Sigma E_{i+2}^2 = 2(1 + R_2 - 4uR_1 + 2u^2) \qquad 2.6$$

This expression has a minimum value when

$$u = \frac{-(-4R_1)}{2 \times 2} = R_1 \qquad 2.7$$

or

$$\cos 2\pi hf = R_1$$

$$f = \frac{1}{2\pi h}[\cos^{-1} R_1 \pm 2n\pi] \qquad 2.8$$

Hence we have the least squares fit for a cosine operator predicting an aribtrary stationary series. We find that f is determined only by the first lag of the auto-correlation function of the series, and that f is only determined modulo $1/h$. This last fact is apparent if we refer back to equation 2.1, where we see that cosine operators have identical forms for angular frequencies differing by $1/h$.

Hence

$$\text{Min} \sum_{i} E_{i+2}^2 = 2(1 + R_2 - 2R_1^2) \qquad 2.9$$

If we want a perfect least squares fit we have

$$R_1 = \pm\sqrt{\tfrac{1}{2}(1 + R_2)} \qquad 2.10$$

with the restriction cosine $2\pi hf = R_1$

One way to meet this condition is to let the interval h shrink toward zero so that $R_1 \rightarrow 1$ and $R_2 \rightarrow 1$. This is equivalent to saying that any small segment of the original series approaches a straight line, in the case where the function is continous and its first derivative exists.

<u>The Geometry of Cosine Operators</u>

A. Error Sum as Function of W

Consider the sum of squared errors as a function of u. We have

$$\sum_{i} E_{i+2}^2 = 2(1 + R_2 - 4uR_1 + 2u^2) \qquad 2.6$$

This is a parabola in u as shown in Fig. 2.1. $u = \cos 2\pi hf$ must lie in the range $-1 \leq u \leq 1$. We have shown that $u = R_1$ is the condition for a minimum fit, and since $-1 < R_1 < +1$, ΣE^2 will always have its minimum in this range.

This means that, for any series, we can always get a minimum fit with some cosine operator of frequency f, where f must be in the range

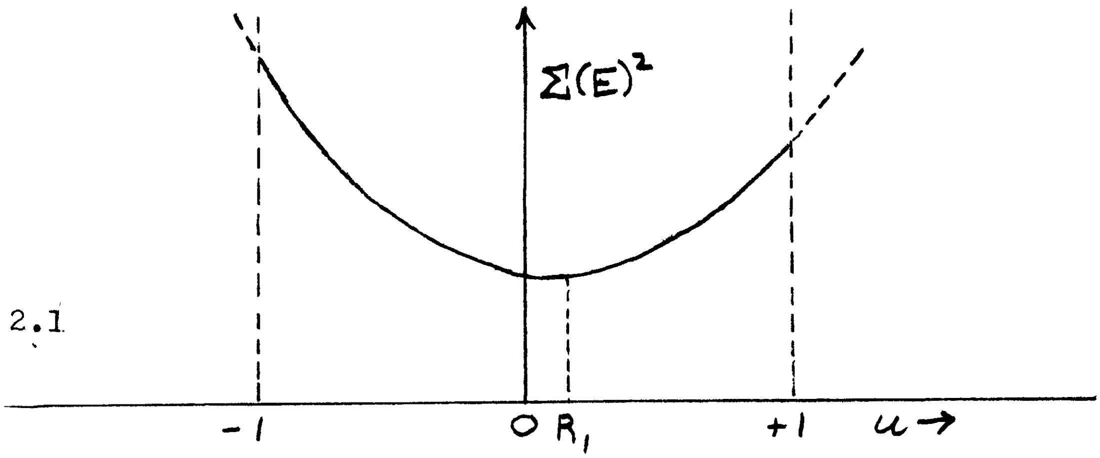


Fig. 2.1

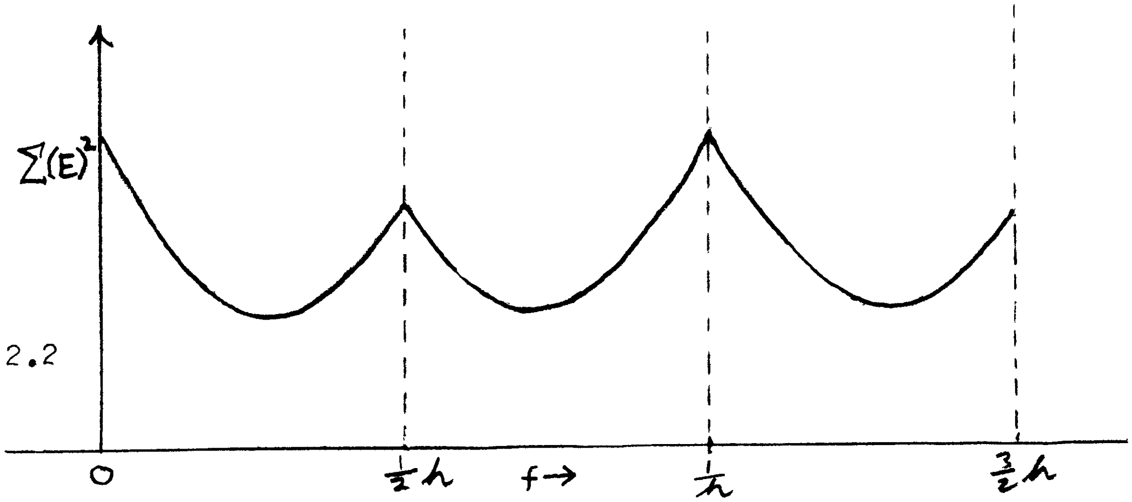


Fig. 2.2

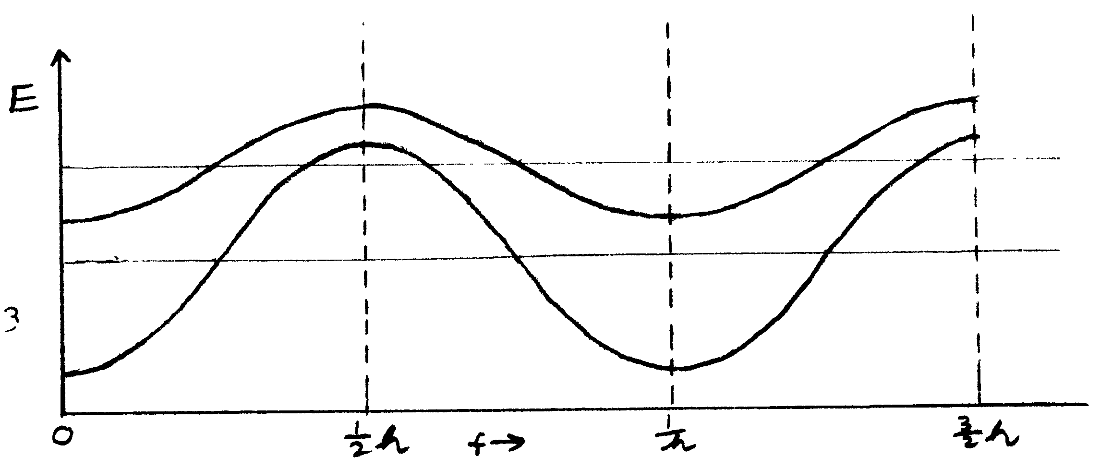


Fig. 2.3

$$0 < f < \frac{1}{2h}$$

$$\begin{pmatrix} u = 1 \\ 2\pi hf = 0 \end{pmatrix} \quad \begin{pmatrix} u = -1 \\ 2\pi hf = \pi \end{pmatrix} \qquad 2.11$$

We require ΣE^2 to be non-negative. This means the discriminant of 2.9 must be ≤ 0 or

$$16 R_1^2 - 8(1+R_2) \leq 0 \qquad 2.12$$

Therefore the curve cannot cross the u axis, but can be tangent to it at one point, when the equality sign holds above. This is the condition for a perfect fit.

B. Error Sum as Function of f

The error sum as a function of frequency f is not truly parabolic but has the general shape of a parabola. It is periodic in f with a period 1/h. It appears as shown in Fig. 2.2 .

C. Individual Errors as Functions of u and f

Equation 2.3 gives us an expression for the individual errors

$$E_{i+2} = X_{i+2} + X_i - 2uX_{i+1} \qquad 2.3$$

If we fix attention on a single individual error (i constant) and let u vary we see that $E\{f\}_{i+2}$ is sinusoidal since $u = \cos 2\pi hf$. Thus $E\{f\}_{i+2}$ varies sinusoidally about a mean given by the sum of the $i^{\underline{th}}$ and the $(i-2)^{\underline{nd}}$ value of the series, with an amplitude of twice the $(i-1)^{\underline{rst}}$ value of the series. The period is 1/h. There is no phase shift between these curves for different i values. Thus all individual errors must increase or decrease simultaneously with f.

Fig. 2.3 shows individual errors as functions of f. This figure explains why the error sum of Fig. 2.2 is reflected across the line f = 1/2h. This line in Fig. 2.3 is the axis of symmetry for the individual errors, so that it must also be the symmetry axis for the error sum.

Conclusions

From the above, we can draw certain conclusions.

1. If one limits himself to the general class of cosine operators, there is a maximum error obtainable for the particular data, using any frequency whatever. That is, there is such a thing as a worst fit for cosine operators.

2. Since ΣE^2 is parabolic, determining 3 values of ΣE^2 is sufficient to determine the complete shape of the error curve for all other frequencies.

3. Moreover, since the individual errors are sinusoidal in f, determining the individual errors for 3 values of f determines the errors for all f.

Looking at the problem another way, much of the information obtainable from any data series by a study of this type is contained in the first and second lags of the auto-correlation function of the series, for these two quantities determine the shape and position of the curve ΣE^2 .

Example

The prediction program described in Appendix B provided a means for testing the conclusions reached about cosine operators. Individual errors and sums of squared errors were computed for cosine operators of frequencies 25, 30, ... 75 cps. The data for which these were computed were readings taken from a typical seismic trace at intervals of 2 ms.

The sums of squared errors are plotted in Fig. 2.4 over two intervals of 240 readings each. Both curves exhibit very good parabolic shapes. The average minimum for the two curves occurs for u = .85 . This should equal the first lag auto-correlation over the two intervals, which was computed by the correlation program (Appendix D) to be .853 .

Fig. 2.4 shows several individual errors plotted as functions of the frequency of the cosine operator used. They appear to be sections of sinusoids as expected.

These curves, computed on an arbitrary time series, seem to be in remarkable agreement with the theory.

Fig. 2.4

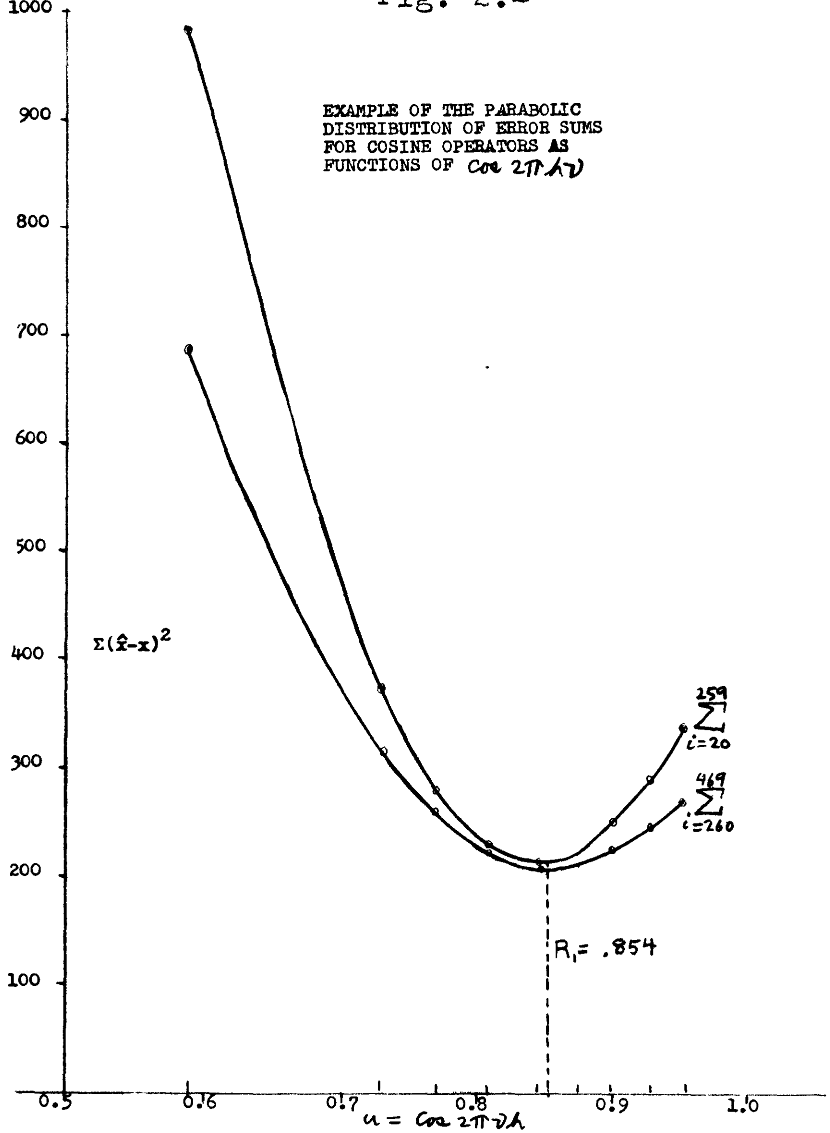
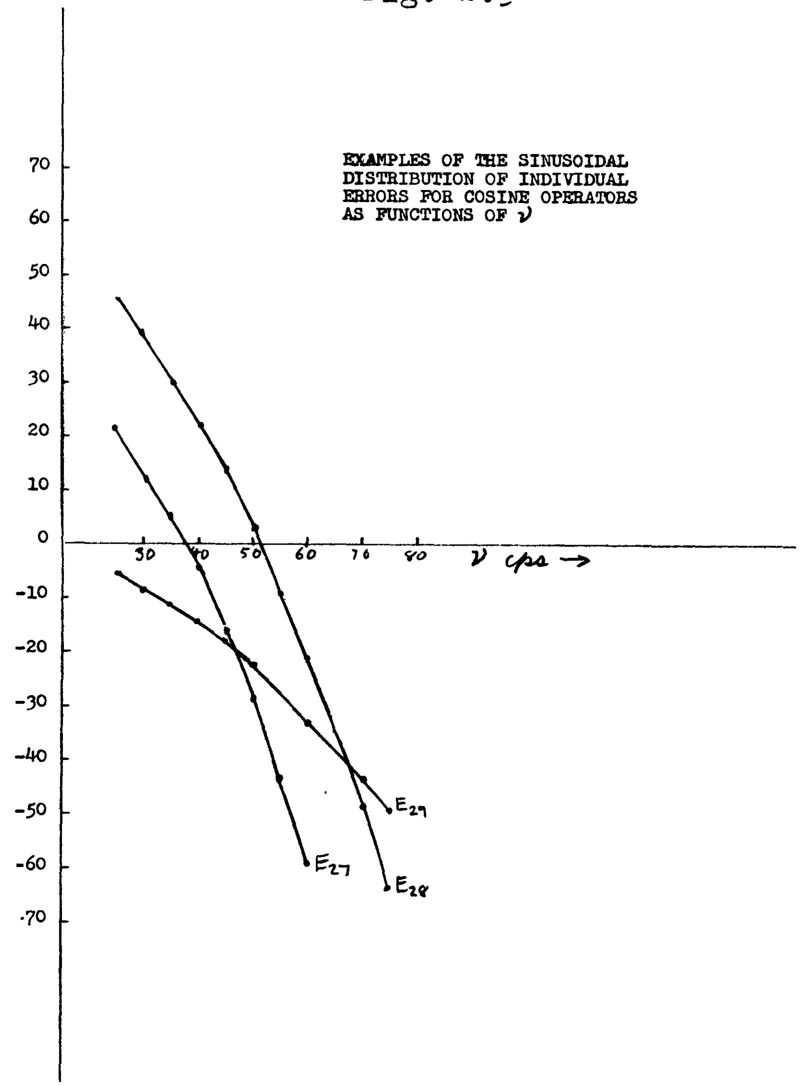

EXAMPLE OF THE PARABOLIC DISTRIBUTION OF ERROR SUMS FOR COSINE OPERATORS AS FUNCTIONS OF Cos 2π h ν
1000
900
800
700
600
500
400
300
200
100
Σ(x̂-x)²
259 Σ i=20
469 Σ i=260
R₁= .854
0.5
0.6
0.7
0.8
0.9
1.0
u = Cos 2π ν h

Fig. 2.5


EXAMPLES OF THE SINUSOIDAL DISTRIBUTION OF INDIVIDUAL ERRORS FOR COSINE OPERATORS AS FUNCTIONS OF ν
70
60
50
40
30
20
10
0
-10
-20
-30
-40
-50
-60
-70
30
40
50
60
70
80
ν cps →
E29
E27
E28

The Cyclical Nature of Cosine Operators

We have mentioned that cosine operators differing in frequency by n/h have identical forms. It is interesting to see what this means physically.

Suppose we are trying to represent a cosine wave of 1 cycle / second with a spacing of $h = 1/4$ second. The points we would plot might appear as in Fig. 2.6 .

Now consider a cosine wave of frequency $1 + 1/h =$ 5 cycles / sec. If we try to plot this frequency with a spacing of 1/4 sec. we find that it can be exactly represented by the points we plotted for the one cycle wave. This is illustrated in Fig. 2.7 . We would find the same would be true for frequencies of $1 + n/h = 1, 5, 9, 13, 17 \ldots.$ Thus, it is the fact that we cannot uniquely represent frequencies differing by n/h that explains the identity of form for cosine operators whose frequencies differ by this amount.

This is also the explanation for the so-called "condensed" power spectra met with in computational procedures.‡ The computed power at a frequency f must represent the sum of the powers at frequencies f, $f, + 1/h$, $f + 2/h \ldots.$ Therefore power spectra can only have the range 0 to $1/h$ cycles. In practice h must be chosen so that $1/h$ is greater than the greatest frequency from which significant contribution is expected.

‡ Ref. 3

Fig. 2.6

COSINE WAVE 1 CYCLE/SEC.

PLOTTING INTERVAL h = $\frac{1}{4}$ SEC.

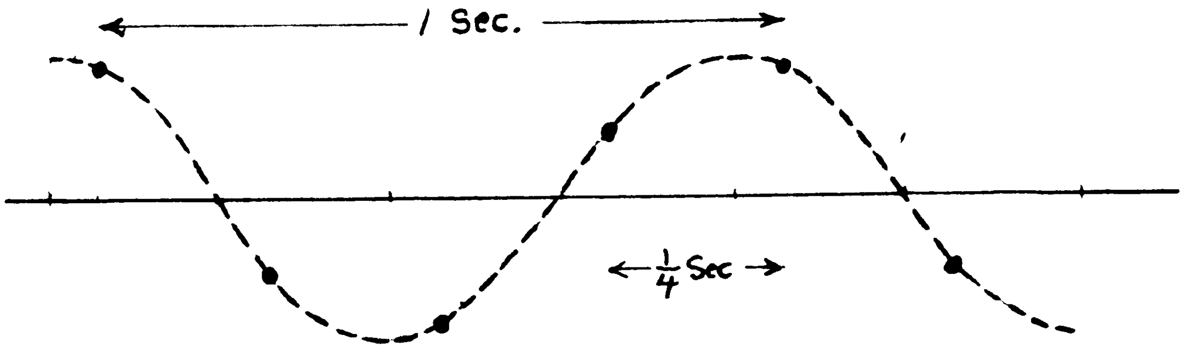
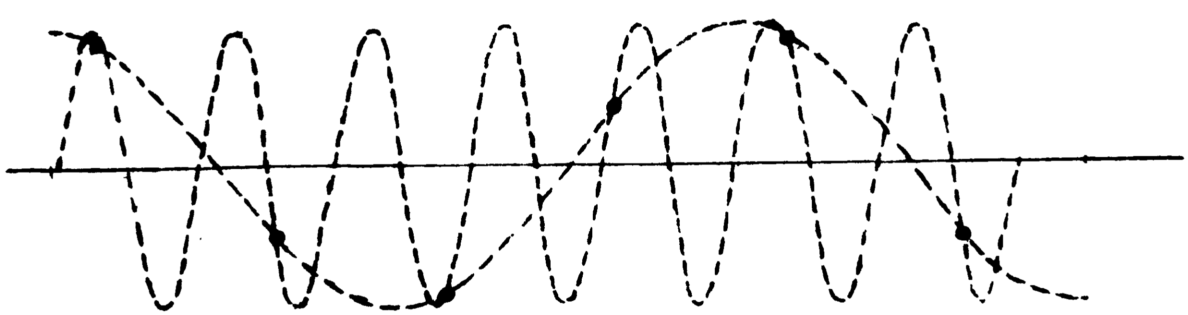


Fig. 2.7

COSINE WAVES 1 AND 5 CYCLES/SEC.



Cosine Operator Predicting an Autoregressive Series

In order to examine the filter characteristics of cosine operators it is convenient to consider their effect on autoregressive-type series. The autoregressive series has a known Cauchy-type distribution of its spectrum.‡ It will be interesting to examine what the spectrum of the error function will be when we predict such a series with a cosine operator.

Referring back to equation 2.3 we have the error function for cosine operators.

$$E_{i+2} = X_{i+2} + X_i - 2uX_{i+1} \qquad 2.3$$

To get the spectrum of this function we first find the auto-correlation R_τ.

$$R_\tau = \Sigma(X_{i+2} + X_i - 2uX_{i+1})(X_{i+2-\tau} + X_{i-\tau} - 2uX_{i+1-\tau}$$

$$R_\tau = \Sigma X_{i+2}X_{i+2-\tau} + \Sigma X_{i+2}X_{i-\tau} + \Sigma X_i X_{i+2-\tau}$$

$$+ \Sigma X_i X_{i-\tau} - 2u(\Sigma X_{i+1}X_{i+2-\tau} + \Sigma X_{i+1}X_{i-\tau}$$

$$+ \Sigma X_{i+1-\tau}X_{i+2} + \Sigma X_{i+1-\tau}X_i) + 4u^2 \Sigma X_{i+1}X_{i+1-\tau}$$

If the X series is properly normalized we may write this in terms of the correlations r_τ of the X series.

$$R_\tau = r_\tau + r_{\tau+2} + r_{\tau-2} + 4u^2 r_\tau$$

$$+ r_\tau - 2u(r_{\tau-1} + r_{\tau+1} + r_{\tau+1} + r_{\tau-1})$$

‡ Ref. 7

$$R_\tau = r_{\tau-2} + r_{\tau-1}(-4u) + r_\tau(2+4u^2) + r_{\tau+1}(-4u) + r_{\tau+2} \qquad 2.13$$

Since the series is taken to be autoregressive

$$r_\tau = \cos 2\pi f_0 \tau h \, e^{-\alpha\tau h} \equiv \cos a\tau \, e^{-b\tau} \quad ‡ \quad \tau \geq 0 \qquad 2.14$$

Substituting equation 2.14 into 2.13 we have

$$\begin{aligned} R_\tau = {} & \cos[a(\tau-2)]e^{-b(\tau-2)} \\ & - 4u\cos[a(\tau-1)]e^{-b(\tau-1)} \\ & + (2+4u^2)\cos a\tau e^{-b\tau} \\ & - 4u\cos[a(\tau+1)]e^{-b(\tau+1)} \\ & + \cos[a(\tau+2)]e^{-b(\tau+2)} \end{aligned}$$

Using trigonometric identities

$$\begin{aligned} R_\tau = {} & e^{2b}[\cos a\tau\cos 2a + \sin a\tau\sin 2a]e^{-b\tau} \\ & - 4ue^{b}[\cos a\tau\cos a + \sin a\tau\sin a]e^{-b\tau} \\ & + (2+4u^2)(\cos a\tau)e^{-b\tau} \\ & - 4ue^{-b}[\cos a\tau\cos a - \sin a\tau\sin a]e^{-b\tau} \\ & + e^{-2b}[\cos a\tau\cos 2a - \sin a\tau\sin 2a]e^{-b\tau} \end{aligned}$$

or

$$R_\tau = \cos a\tau e^{-b\tau}[e^{2b}\cos 2a - 4ue^{b}\cos a + 2 + 4u^2 - 4ue^{-b}\cos a + e^{-2b}\cos 2a]$$

‡ Ref. 5

$$+ \quad \sin a\tau e^{-b\tau}[e^{2b}\sin 2a - 4ue^{b}\sin a$$

$$+ 4ue^{-b}\sin a - e^{-2b}\sin 2a]$$

Hence

$$R_\tau = A\cos a\tau e^{-b\tau} + B\sin a\tau e^{-b\tau}$$

$$= e^{-b\tau}(A\cos a\tau + B\sin a\tau) \qquad 2.15$$

where

$$A = \cos 2a(e^{2b}+e^{-2b}) - 4u\cos a(e^{b}+e^{-b}) + 2 + 4u^2$$

$$B = \sin 2a(e^{2b} - e^{-2b}) - 4u\sin a(e^{b} - e^{-b}) \qquad 2.16$$

$$R_\tau = e^{-b\tau}(A^2+B^2)^{1/2}\left[\frac{A}{(A^2+B^2)^{1/2}}\cos a\tau + \frac{B}{(A^2+B^2)^{1/2}}\sin a\tau\right]$$

$$= e^{-b\tau}(A^2+B^2)^{1/2}(\cos a\tau\cos\beta + \sin a\tau\sin\beta)$$

$$\text{where } \beta = \tan^{-1}\frac{B}{A}$$

Thus

$$R_\tau = e^{-b\tau}(A^2+B^2)^{1/2}\cos(a\tau+\beta) \qquad 2.17$$

R_τ is now in a form similar to the r_τ for the original series. The spectrum of this type is known to be a Cauchy distribution. The specific shape will be controlled by values of b, A, B, and β.

Rather than continuing with this example we shall proceed to another type of series. The autoregressive series is somewhat non-typical. Its spectrum, the Cauchy distribution, is very broad, in fact there is no mean value of frequency for this spectrum.

Cosine Operators on Series with Gaussian Spectrum Distribution

A much more stringent series than the autoregressive type is a series with a power spectrum composed of two Gaussian curves. The spectrum has the form ‡

$$\phi(\omega) = \frac{1}{2\sigma\sqrt{2\pi}}\left[e^{\frac{-(\omega-a)^2}{\sigma^2}} + e^{\frac{-(\omega+a)^2}{\sigma^2}}\right] \qquad 2.18$$

where +a and -a are the respective means of the two Gaussian curves, and σ is their standard deviation in radians. With such a series the normalized auto-correlation function may be written as ‡

$$r_\tau = e^{-\frac{\sigma^2\tau^2}{2}} \cos a$$

If we predict such a series with a cosine operator, we generate an error series whose auto-correlation function is, as before

$$R_\tau = r_{\tau-2} + r_{\tau-1}(-4u) + r_\tau(2+4u^2) + r_{\tau+1}(-4u) + r_{\tau+2} \qquad 2.13$$

or substituting

$$\begin{aligned} R_\tau = \; & e^{-\frac{\sigma^2(\tau-2)^2}{2}} \cos[a(\tau-2)] \\ & + e^{-\frac{\sigma^2(\tau-1)^2}{2}} \cos[a(\tau-1)]\,(-4u) \\ & + e^{-\frac{\sigma^2(\tau)^2}{2}} \cos[a\tau]\,(2+4u^2) \\ & + e^{-\frac{\sigma^2(\tau+1)^2}{2}} \cos[a(\tau+1)]\,(-4u) \\ & + e^{-\frac{\sigma^2(\tau+2)^2}{2}} \cos[a(\tau+2)] \end{aligned} \qquad 2.20$$

‡ Ref. 5

This can be reduced, as before, to the form

$$R_\tau = e^{\frac{-\sigma^2\tau^2}{2}} [A(\tau)\cos a\tau + B(\tau)\sin a\tau] \qquad 2.21$$

where

$$A(\tau) = \cos 2a \left[e^{-\frac{\sigma^2(-4\tau+4)}{2}} + e^{-\frac{\sigma^2(4\tau+4)}{2}}\right] - 4u\cos a\left[e^{-\frac{\sigma^2(-2\tau+1)}{2}} + e^{-\frac{\sigma^2(2\tau+1)}{2}}\right] + (2 + 4u^2)$$

$$B(\tau) = \sin 2a \left[e^{-\frac{\sigma^2(-4\tau+4)}{2}} + e^{-\frac{\sigma^2(4\tau+4)}{2}}\right] - 4u\sin a\left[e^{-\frac{\sigma^2(-2\tau+1)}{2}} + e^{-\frac{\sigma^2(2\tau+1)}{2}}\right] \qquad 2.22$$

If we are interested in the power spectrum of this series we want

$$\phi(\omega) = 2\int_0^\infty R(\tau)\cos\omega\tau\, d\tau \qquad 2.23$$

Probably this integral cannot be expressed in closed form, and we shall have to resort to a computed example.

Computational Example

Here we illustrate the filter characteristics of cosine operators in a particular case. We choose a series with a Gaussian spectrum peaked at 50 cycles and with a standard deviation of 22.36 cycles. The power spectrum of such a series is shown in Fig. 2.8 , and was computed from equation 2.18 . In general shape this is not unlike power spectra dealt with on seismic traces. Fig. 2.9 shows the normalized auto-correlation function for this type of series, as derived from equation 2.19 . The series has essentially no correlation for lags greater than about .03 sec.

To examine the effectiveness of cosine operators as frequency filtering mechanisms, a cosine operator of frequency 50 cps was taken. The spacing interval was chosen to be 2.5 ms. The auto-correlation function of the error series generated by this operator is shown in Fig. 2.9 , and is computed from equation 2.13 . In this case the function is unnormalized so that the zeroth lag auto-correlation is proportional to the total power contained in the power spectrum of the error series. Thus we see that less than 20 per cent of the power contained in the original Gaussian series remains in the error series. More than 80 per cent has been "filtered" out. However, since some of this is due merely to curve continuity, the shape of the spectrum of errors is more important than the total power.

The unnormalized spectrum of the error series is shown in Fig. 2.8 , and, as might be expected, is definitely bimodal. This curve clearly indicates that the operator is acting as a filter peaked at 50 cps, at which frequency all power has been removed. Lower frequencies are also well reduced but the higher ones are not so much affected. In fact, the power at 100 cycles is slightly greater than in the original series. This is not a computational error. As discussed below it seems to be a necessary characteristic.

A more convenient way of showing the filter characteristics is to plot the quantity

$$\frac{\text{Power removed at } \omega}{\text{Initial power at } \omega}$$

This graph is shown in Fig. 2.10 . It shows how frequencies lower than 50 cycles are much preferred to those greater. It is possible that this curve would not represent the filter characteristics of a 50 cycle cosine operator used on another type of series. There is some reason, however, to suspect that it does, and that, in fact, the curve of Fig. 2.10 continues downward considerably below the axis (thus representing amplification rather than filtration). If we were to use a series containing mostly frequencies between 100 and 200 cycles, the 50 cycle operator would yield very high errors of prediction. The sum of squared errors would be far from the minimum of Fig. 2.1 . Hence the power in the error series would probably be greater than that in the original series. This could only come about by an amplification of certain frequencies, which would naturally occur for frequencies greatly different from 50 cycles. In this example 200 cycles is chosen as an upper limit, because with a spacing of 2.5ms unique curves only exist from 0 cycles to 1/2h or 200 cycles.

Fig. 2.8

POWER SPECTRA

——— For series with Gaussian spectrum with σ = 22.4cps a = 50 cps (normalised).

----- For error series generated by a cosine operator of frequency 50 cps on above series (unnormalized).

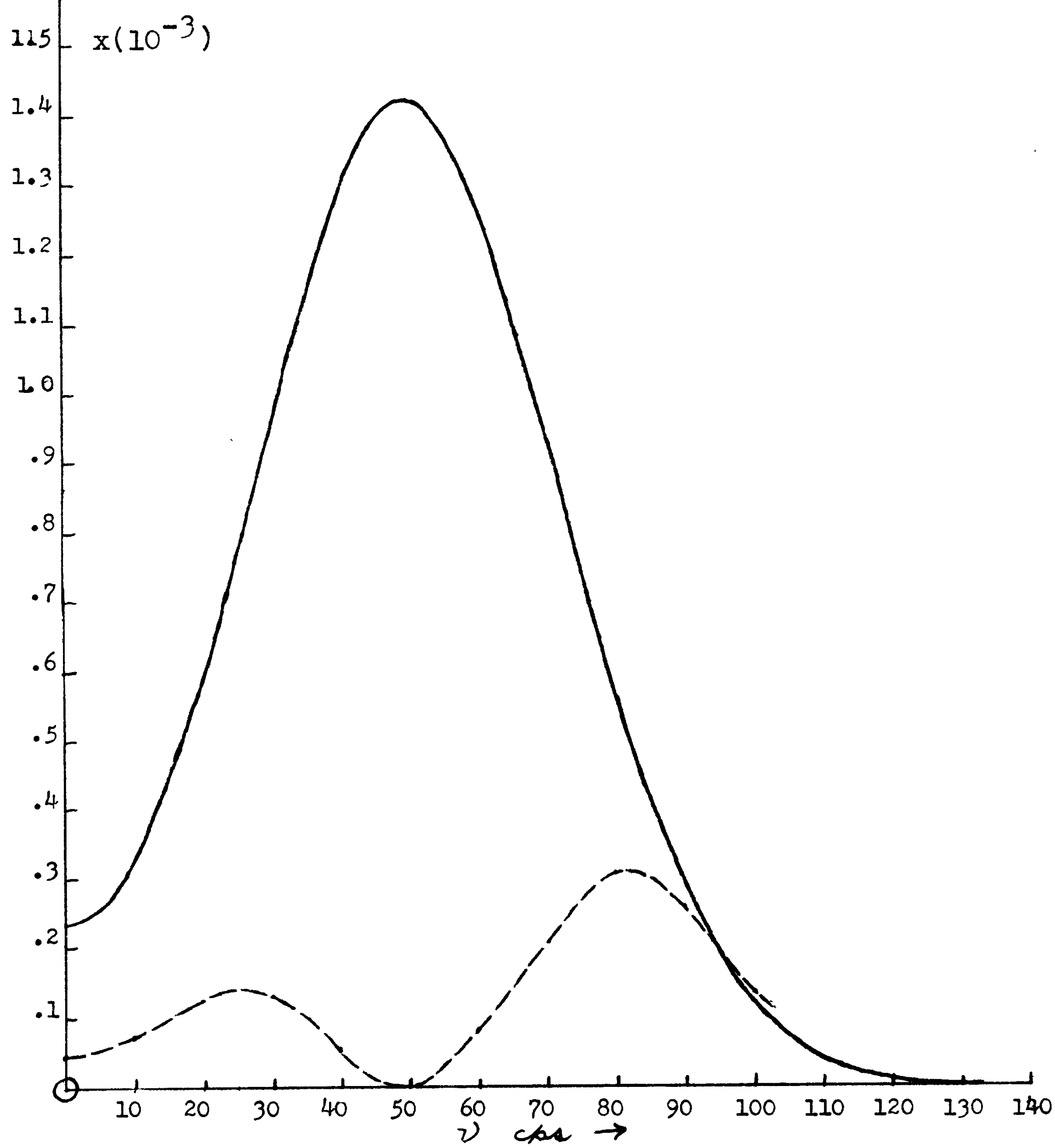
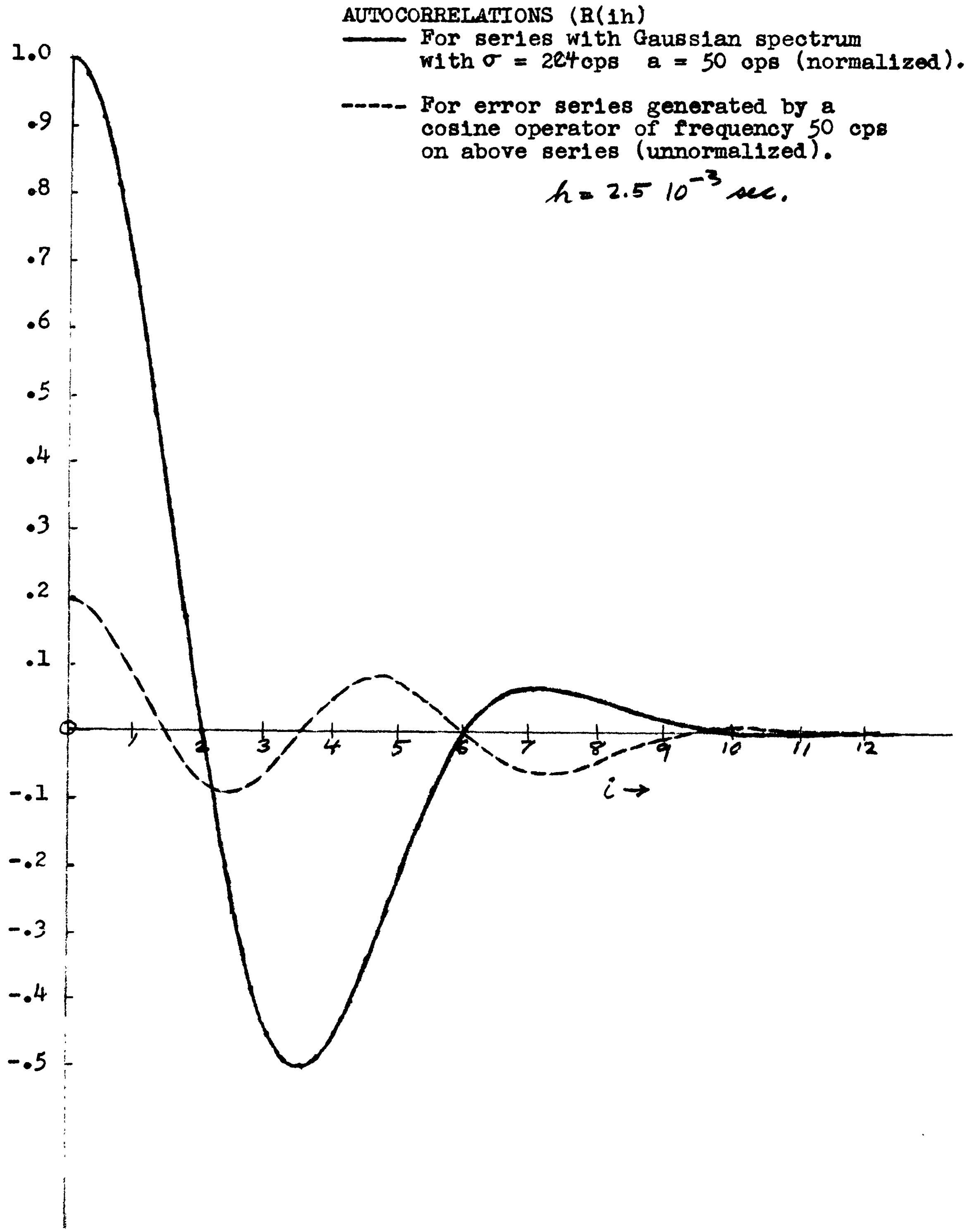


Fig. 2.9



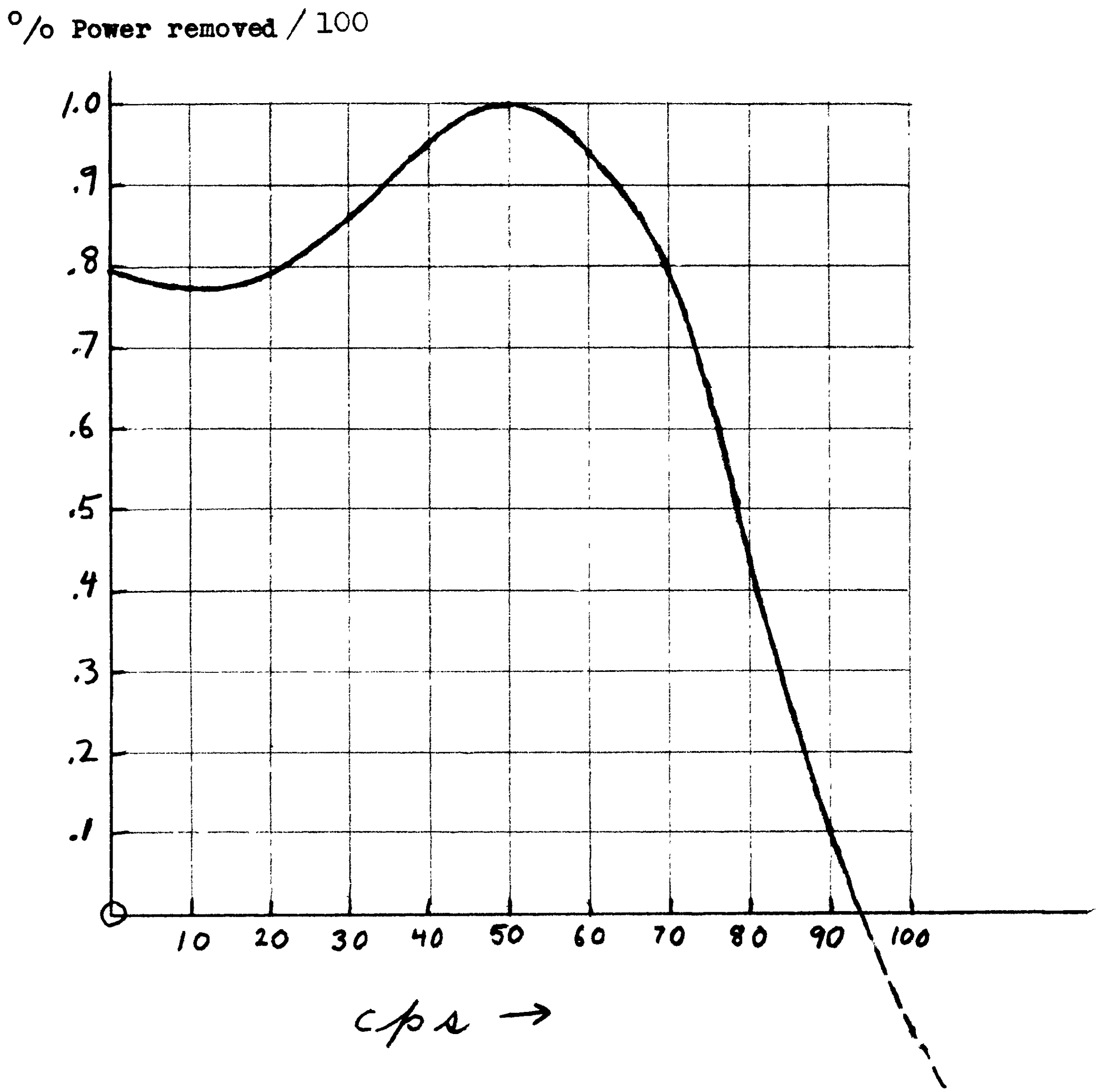
AUTOCORRELATIONS (R(ih)
—— For series with Gaussian spectrum with σ = 224 cps a = 50 cps (normalized).
----- For error series generated by a cosine operator of frequency 50 cps on above series (unnormalized).
h = 2.5 10^-3 sec.
1.0
.9
.8
.7
.6
.5
.4
.3
.2
.1
-.1
-.2
-.3
-.4
-.5
0
1
2
3
4
5
6
7
8
9
10
11
12
i →

Fig. 2.10

FILTER CHARACTERISTICS

FOR 50 CYCLE COSINE OPERATOR

WITH h = 2.5 ms.



A Method of Finding Linear Operators for Least Squares Fitting Procedures

Described in this section is an iterative method of approaching the values of coefficients for a least squares fitting of linear operators to multiple time series. The problem arose in connection with the determination of linear operators to use in picking reflections from seismograms. ‡ The method is extremely inefficient and is really only possible with the aid of very high speed computing machines, but it gives interesting insight into the behaviour of matrices, which helps in constructing other techniques.

We are trying to fit a linear operator of the form

$$\begin{aligned} \hat{x}_{i+k} = c &+ a_0 x_i + a_1 x_{i-1} \dots + a_M x_{i-M} \\ &+ b_0 y_i \quad \dots\dots\dots\dots + b_M y_{i-M} \\ &+ c_0 z_i \quad \dots\dots\dots\dots + c_M z_{i-M} \\ &+ d_0 u_i \quad \dots\dots\dots\dots + d_M u_{i-M} \end{aligned} \qquad 2.24$$

to an interval of the sequences x_i, y_i, z_i, and u_i so that

$$I = \sum_i (x_{i+k} - \hat{x}_{i+k})^2 \text{ is a minimum.} \qquad 2.25$$

The plan is to guess initial values of the constants a, a_s, b_s, c_s, and d_s and compute 2.24 . Then adjust the constants so that I is continually reduced. The initial

‡ Ref. 5

values chosen are $a = \bar{x}$ (mean of x_1 series) $a_s = b_s = c_s = d_s = 0$. These values are the values which the constants would assume under least squares fitting procedures if the x_1 series were truly random and had no predictability. With these values of the constants $I = I(\bar{x},0,0, \ldots)$ becomes the sample variance about the sample mean.

The computational procedure is:

1. Find $I(\bar{x},0,0, \ldots)$
2. Find $I(\bar{x}+ \Delta a, o,o, \ldots)$
3. If $2 < 1$, continue adding Δa until $I(\bar{x}+n \Delta a, 0,0, \ldots) > I(\bar{x}+(n-1) \Delta a, 0,0, \ldots)$.
   If $2 > 1$, subtract Δa and continue to subtract until $I(\bar{x}-n \Delta a, 0,0, \ldots) >$ $I(\bar{x}-(n-1) \Delta a, 0,0, \ldots)$
4. Using $\bar{x} \pm (n-1) \Delta a, 0,0, \ldots$, as the new starting point, find $I(\bar{x} \pm (n-1) \Delta a, \Delta a_0, 0,0, \ldots)$ and repeat the steps under 3.
5. Work successively in this fashion with each of the variables a, a_0, $a_1 \ldots d_M$.
6. Start the process over again with the variable a.
7. Continue recycling until the desired accuracy is reached.

It is interesting to consider the geometry of this process. If we substitute equation 2.24 into 2.25, we find that I is parabolic in each of the coefficients a, a_s, b_s, c_s, d_s. For simplicity consider the case where we have only two coefficients a and b. Then I is a two dimensional paraboloid in a and b whose minimum we wish to find. I is positive or zero for all a, b and has one minimum. Contours of I = c are ellipses in the a b plane of constant major to minor axis ratios, and are centered at the minimum. Figs. 2.11, 2.12, and 2.13 illustrate three situations that might arise. In Fig. 2.11 the contours are circular which is the case when the matrix of the normal equations associated with the minimum fit is well-behaved. Fig. 2.12 is the more usual situation where the contours are definitely elliptial. Fig. 2.13 shows a very badly-behaved situation corresponding to near singularity of the associated matrix.

The solid line shows how the iterative method described above would converge toward the minimum point in the three situations. The dashed curve shows how another iterative method, the steepest descent method, would converge in these situations. The steepest descent method runs into trouble in the near singular case because with finite increments it cannot land on the long axis of the ellipse. It is forced to wobble back and forth, much as a small ball would wobble rolling in such a trough. The method described above would also encounter trouble if the increment were not fine enough, for if it got near the

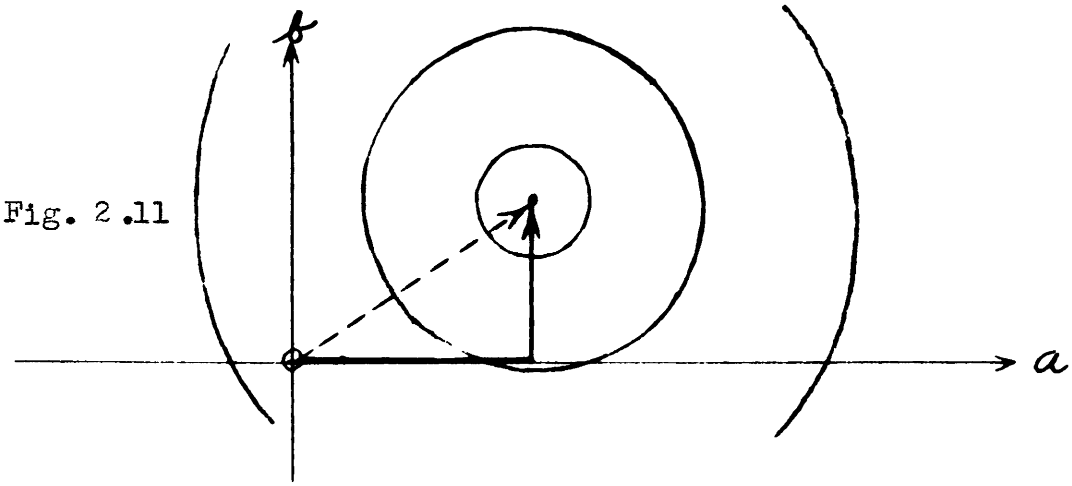


Fig. 2.11

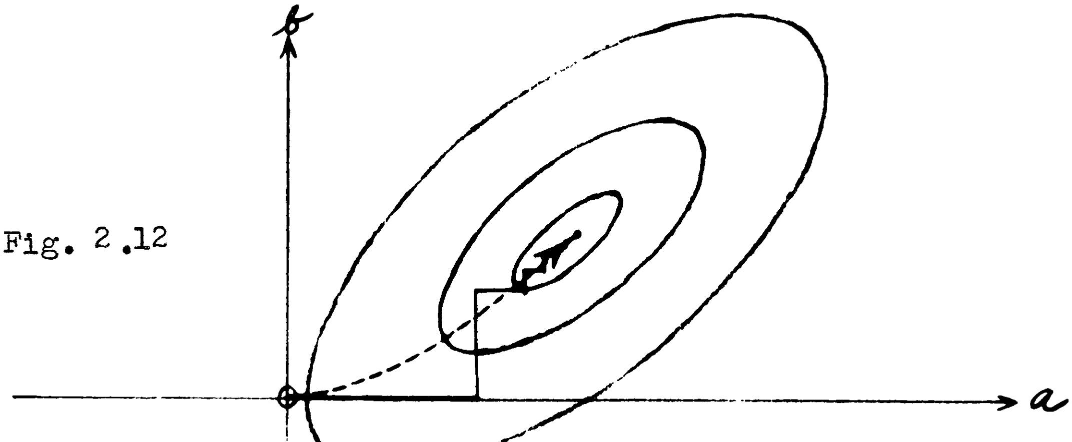


Fig. 2.12

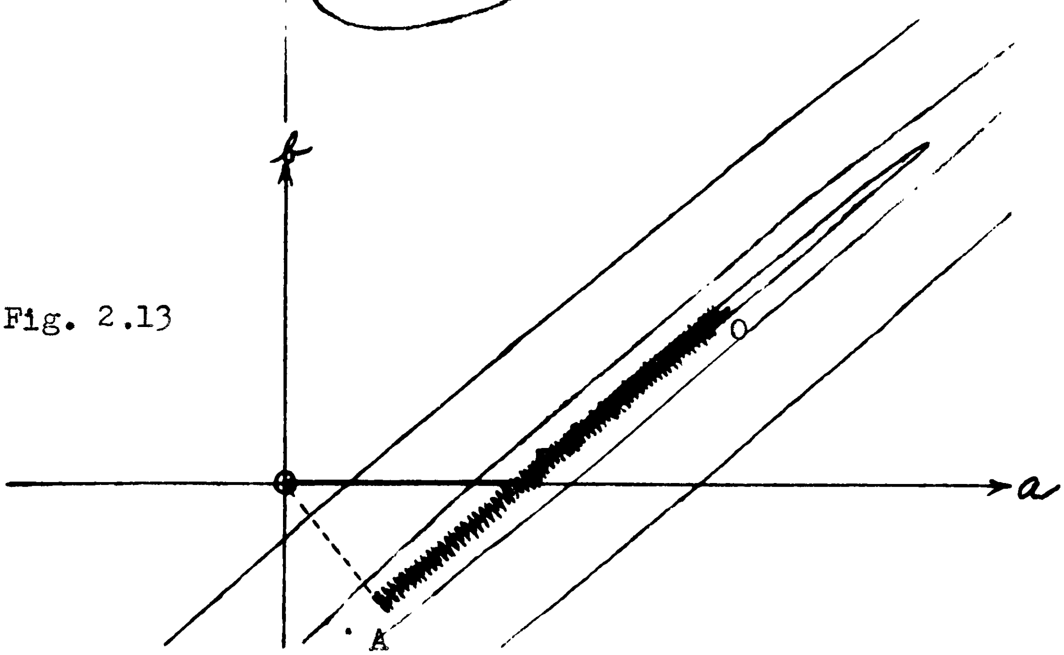
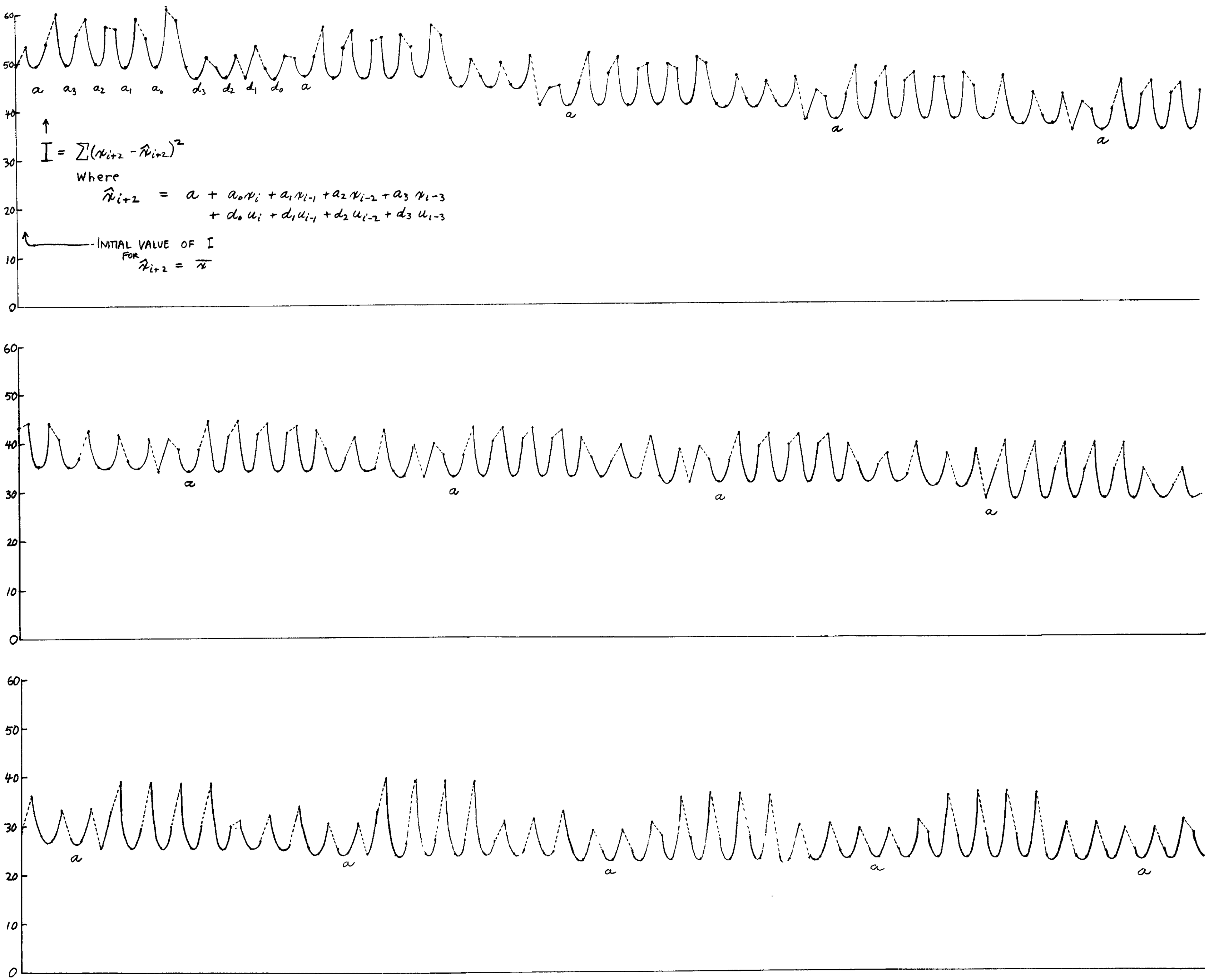


Fig. 2.13

trough the next increment would carry it across the trough to a greater value of I.

These figures illustrate a fundamental problem met with in iterative methods. The fine increments necessary in treating the near singular case are very inefficient when used on well-behaved data, whereas the larger increments applicable in Fig. 2.11 could never find the minimum of Fig. 2.13 .

A program was written for the WWI Digital Computer which would do this one-variable-at-a-time type of iteration. It is described in Appendix C. The computations it carries out take fifteen or twenty minutes of machine time, but they represent nearly a year of hand computation. The program can print out each successive value of I as it is computed. Fig. 2.14 shows a plot of these values as the program converges towards the solution of a particular problem. This diagram shows how I is parabolic in each coefficient. We also note that all the parabolas have approximately the same shape. This indicates that if there is a predominant long ellipse axis as in Fig. 2.13 , it cannot be close to parallel to any of the axes a, a_s, d_s, for if it were, the parabolic section in the corresponding direction would be quite flat. One surprising feature of this diagram is the failure of the parabolas to tend to flatten as I is diminished.



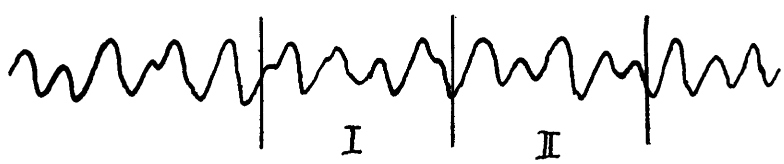
60
50
40
30
20
10
0
a a3 a2 a1 a0 d3 d2 d1 d0 a
I = ∑(x_{i+2} - x̂_{i+2})^2
Where
x̂_{i+2} = a + a0 x_i + a1 x_{i-1} + a2 x_{i-2} + a3 x_{i-3} + d0 u_i + d1 u_{i-1} + d2 u_{i-2} + d3 u_{i-3}
INITIAL VALUE OF I FOR x̂_{i+2} = x̄

Fig. 2.14

Accuracy

This is a convenient point to consider the problem of the importance of obtaining the exact solution. If we look at Fig. 2.13 , we see that values of a and b at the point A will reduce I almost as well as values at the true minimum O. Individual errors $(x_1-\hat{x}_1)$ will likewise be practically identical. The effect of the displacement OA will not be felt until the values at A are used to predict outside the interval where the minimum fit is taken.

Suppose the series is



and the minimum fit is taken in the interval I of this series. What happens when we predict the interval II with coefficients chosen in I?

Consider Fig. 2.15 . The dark solid line represents the long axis of the ellipses for interval I and the light solid lines, the contours for this interval. The true minimum of these contours is at O. Likewise, we can draw a similar contour picture for the interval II. If we assume the dynamics are but slightly different in the two intervals, the second contours will be slightly rotated with respect to the first, and, there will be a small displacement of the minimum. The heavy and light dashed lines in Fig. 2.15

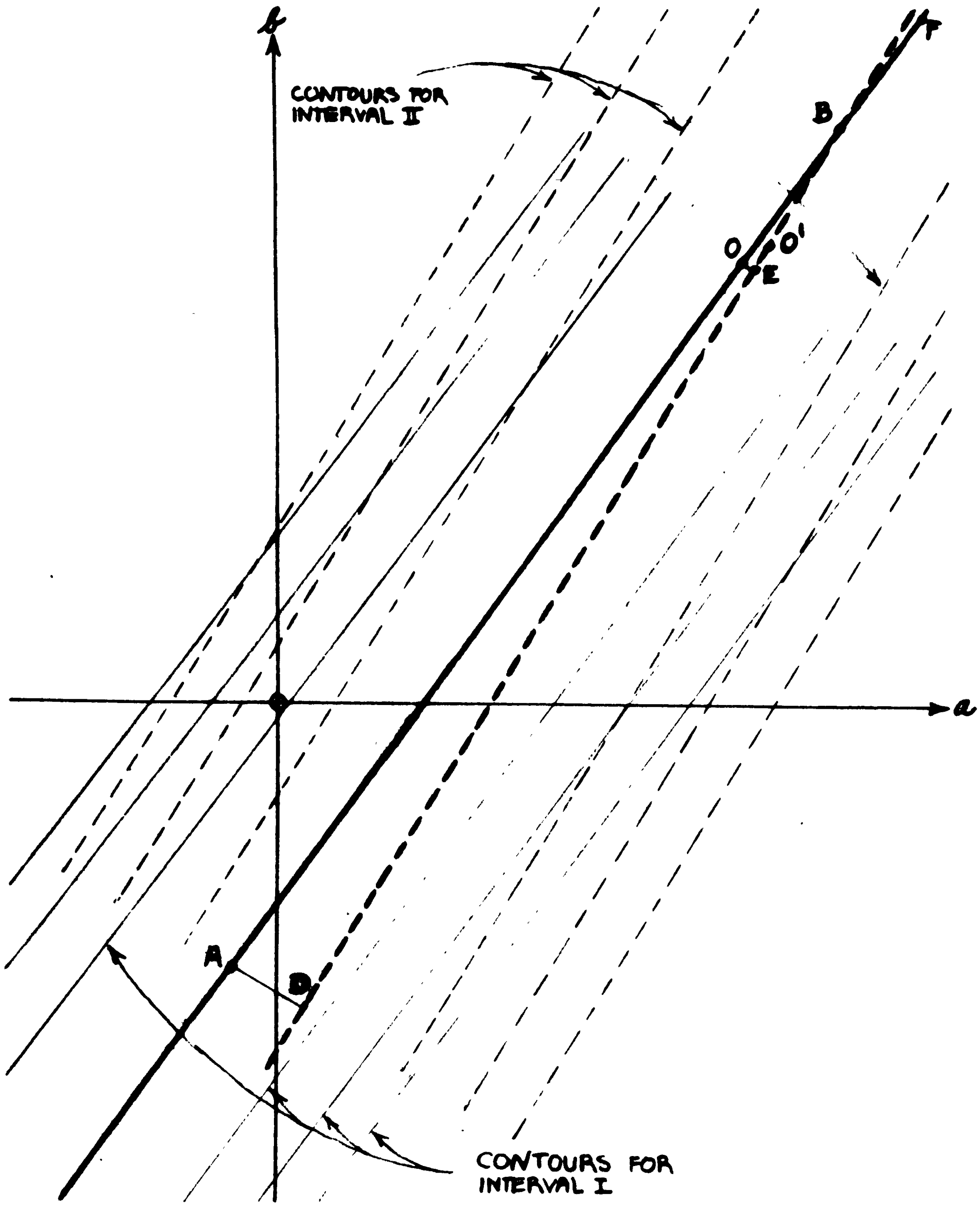

b
CONTOURS FOR
INTERVAL II
B
F
O
O'
E
a
A
D
CONTOURS FOR
INTERVAL I

Fig. 2.15

represent these contours for interval II. Since we are considering the near singular case, the deviation of the sum of squared errors for the second interval from its value on the heavy dashed line will vary as the square of the distance from a point in the ab plane to the heavy dashed line.

Now suppose in finding our minimum point for interval I we had landed at the point A which satisfies the least squares criterion almost as well as the true point O. The deviation of the sum of squared errors when point A is used to predict interval II will be proportional to $(AD)^2$ which would be about sixteen times greater than if the point O were used, since OE - 1/4 AD. On the other hand, if we had landed at the point B for the first interval we would get a sum of squared errors smaller than if the point O were chosen. Again, if the point F were taken, the sum of squared errors for interval II would not be appreciably different than for the point O.

These effects have been noted in computed data. The indication is that the true minimum point O must be chosen if we are to take the sum of squared errors as a valid comparison of the changing dynamics in various intervals of a series by this method.

PART III

SOME INTERPRETIVE PROCEDURES

Introduction

In this part we present several ideas which may be applicable to answering certain questions involving seismogram analysis. Two of the ideas have had some testing, the others none. With one exception these ideas relate specifically to reflection seismic records, and various possibilities in picking reflections therefrom. The questions are:

1. In a two velocity system, e.g., shear and compressional waves, can we set up a method for separating these velocities and can we apply it to reflection determination?
2. In the use of linear operators for seismogram analysis, is there another measure of prediction error, other than the "error curve", which will show reflections?
3. Can we obtain information on the step-out times of reflections, by the use of linear operators and the concept of ensemble averages?
4. Can a special seismometer set-up be used in conjunction with correlation analysis to pick reflections?

Velocity Separation

The determination of velocities for compressional waves in the earth at shallow depths is relatively simple due to (1) the ease in generating such waves, and (2) the fact that the first arrivals are the compressional waves. Shear waves are more difficult to generate with sifficient amplitude to separate from the earlier arriving types. Although with the proper equipment this can be done by visual inspection of the seismogram, ‡ it seemed of interest to consider if a statistical test could be devised to help in this problem.

The approach was to set up a simple model approximating the physical situation.

Assume we have two wave forms A and B traveling horizontally at velocities V_A and V_B, where V_A V_b, past three geophones F, G, and H, equally spaced with separation d. The wave shapes do not change with time. Traces F, G, and H then represent composites of A and B with different time lags. Assuming V_A is known, the problem is to find V_B and, if possible, the wave forms A and B.

Divide the time scale into units such that the no. of units per sec. is L. Since V_A and d are known we may line up F, G, and H so that very nearly

‡ Ref. 8

$$F_N = A_N + B_N \qquad 3.1$$

$$G_N = A_N + B_{N-j} \qquad 3.2$$

$$H_N = A_N + B_{N-2j} \qquad 3.3$$

where the time lag between traces is approximated by j units so that $\frac{j}{L} \doteq \frac{-d}{V_A} + \frac{d}{V_B}$

or

$$V_B = \frac{LdVA}{jV_A+Ld} \qquad 3.4$$

This is illustrated in Fig. 3.1 .

From equations 3.1, 3.2, and 3.3 we can get

$$A_N - A_{N-j} = G_N - F_{N-j} \qquad 3.5$$

$$A_N - A_{N-2j} = H_N - F_{N-2j} \qquad 3.6$$

3.5 and 3.6 are recursion formulas giving A_{N-kj} and A_{N-2kj} respectively (k = 1, 2, ...) once A_N .... A_{N-j+1} are known. Now if j has its correct value then it is easy to show that <u>regardless of how we choose the initial A's</u> both formulas give the same value for A_{N-2kj} . If j is slightly wrong then the two series will differ slightly. The difference will increase as j strays further from its true value. We may now set up a procedure for finding this value. Assume values for j and for A_N, A_{N-1}, ... A_{n-j+1}, use equations 3.5 and 3.6 to calculate the two series (to

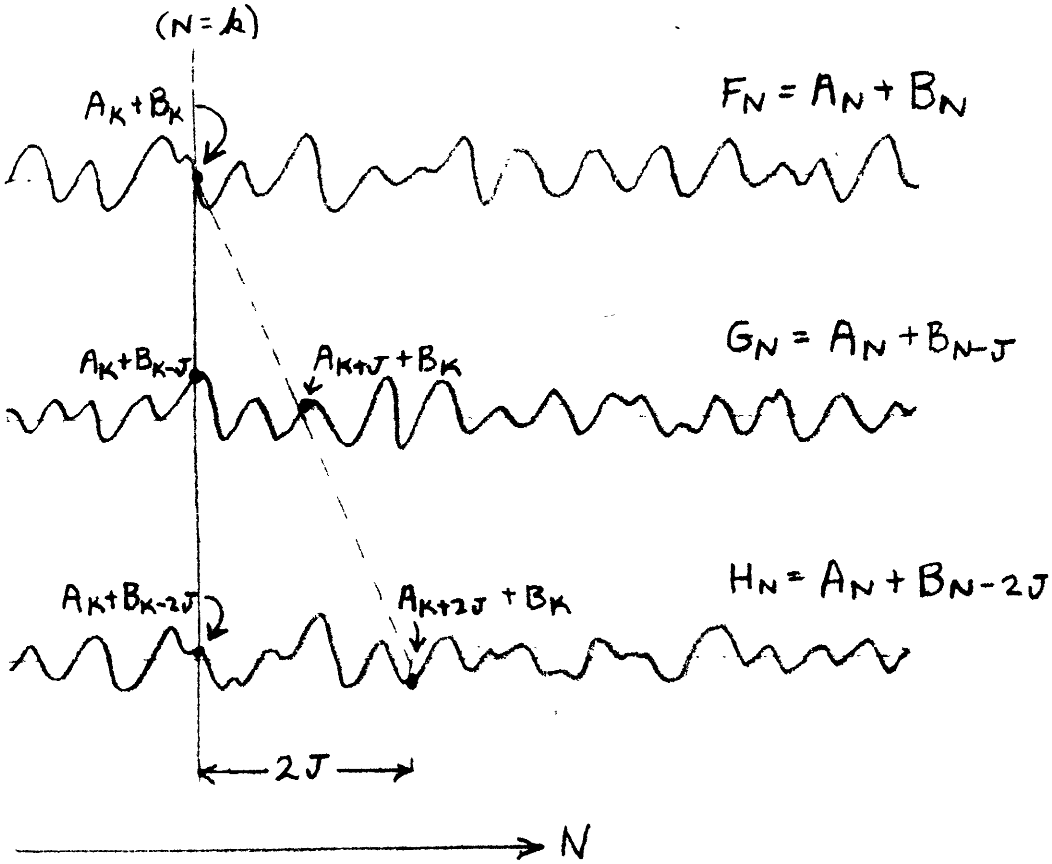

(N = k)
Ak+Bk
FN = AN + BN
Ak+Bk-J
Ak+J +Bk
GN = AN + BN-J
Ak+Bk-2J
Ak+2J +Bk
HN = AN + BN-2J
2J
N

Fig. 3.1

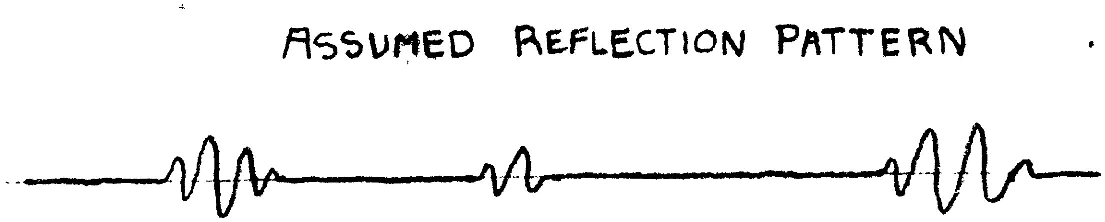

ASSUMED REFLECTION PATTERN

Fig 3.2

a certain length), find the mean square difference between the series, and plot this difference as a function of j. In the ideal case, this function will go to zero for the correct value of j. In practice we can expect this difference function to have a minimum at the correct value.

Thus in theory at least, j is determinable. Equation 3.4 may be used to find V_B. Although the exact wave shapes are indeterminate in the general case, there may be obtained some information about them. Assume the first j values of A_N are taken to be zero. If j is correct, then the series from equation 3.5 represents the true A_N with the first j values subtracted successively. Thus the series from equation 3.5 might be expected to have the same frequency characteristics as the true A_N.

In certain cases the assumption that $A_N \ldots A_{N-j+1} = 0$ will be fairly accurate. In these cases the wave forms should be determinable. Examples would arise in the separation of shear and compressional waves where it is known that the shear waves arrive late, and in reflection picking.

Another possibility in this problem would be the use of pure cross-correlation between two traces. We should expect to get a peak in the correlation at a lag corresponding to the velocity V_B and the particular geophone separation. However, if the wave form B were of small amplitude, the shape of the cross-correlation curve would effectively be dominated by that of the auto-correlation of wave form A, and the selection of the peak would be somewhat arbitrary. On the other hand, the mean square difference between equations 3.5 and 3.6 should still show a true minimum at the correct lag.

We can adapt this idea to the selection of reflections on seismic records. Here we make the simplified assumptions that the reflection consists of a wave train with zero amplitude between reflections as in Fig. 3.2 . This is assumed to occur on two traces in the same form and at the same time (i.e., there is no step out time of the reflection which is assumed to be coming in vertically). In this case equation 3.5 alone is applicable and we need only two traces.

$$A_N - A_{N-j} = G_N - F_{N-j} \qquad 3.5$$

j is taken from the step-out time of the initial breaks on the seismogram. We then select some interval j units in length in which A_N is zero (a non-reflection interval), and use equation 3.5 to predict the remainder of the reflected wave. For interpretation it is convenient to plot the running variance of the predicted reflection.

Now the assumptions will certainly not be upheld exactly on any real seismic record. A certain amount of random energy will be in phase between any two traces and would be picked out by this method as part of the predicted reflection. To alleviate this situation we can use three traces and predict the reflected wave from the three possible pairings of these traces. Adding the three predicted waves would tend to accentuate components in phase between all three, and to minimize the random, in-phase components between any two traces. For three traces F_N, G_N, and H_N we can express this sum as

$$3A_{N+K} = A_N + A_{N+j} + A_{N+K-j} + G_{N+K} + 2H_{N+K} - F_N - F_{N+K-j} - G_{N+j} \qquad 3.7$$

where j corresponds to the step-out between F_N and G_N, and K the step-out between F_N and H_N.

<u>Tests of the Method</u>

1. Selection of Shear Velocity

An initial test was constructed which showed that, when the assumptions were exactly upheld, the minimum of the plot of the squared differences between equations 3.5 and 3.6 was quite sharp.

On this basis three adjacent traces of a seismogram were converted to numerical form and the method applied to these real series. The seismogram was taken at Revere Beach, Mass., in unconsolidated sediments, by Peter Southwick.‡ Special generating apparatus was used so that the shear arrivals were quite prominent. This record is now lost, but

‡ Ref. 8

Fig. 3.3 shows a very similar seismogram taken with the same apparatus. The first line of check marks on this seismogram indicates the first arrivals, and the second line of check marks was picked as the arrivals of the shear waves. This second line permitted a direct computation of the shear velocity.

The readings for the three traces were lined up in accordance the first line of time breaks, and equations 3.5 and 3.6 were computed for a variety of values of j. In each case the first j values of A_N were assumed to be zero. The sum of squared differences between these two series were computed for each j, and normalized by the number of terms in the series for each j. A plot of these quantities appears in Fig. 3.4 .

This figure shows two distinct minima (at j = 13.3 and j = 16.0) rather than just one. Upon examination it turned out that the value j = 13.3 corresponded to a shear velocity which would have been computed by direct interpretation of the first two traces chosen. The second minimum corresponded to a velocity which would have been determined directly from the second and third traces chosen. The value of velocity computed by the <u>entire</u> second line of check marks of Fig. 3.3 lay between these two values.

Fig. 3.5 shows a running average of the points in Fig. 3.4 (by overlapping groups of three) which exhibits a flat minimum between j = 13.3 and j = 16.0 . The

LEXINGTON S+9

Shear test - med. sand

Geophones in line

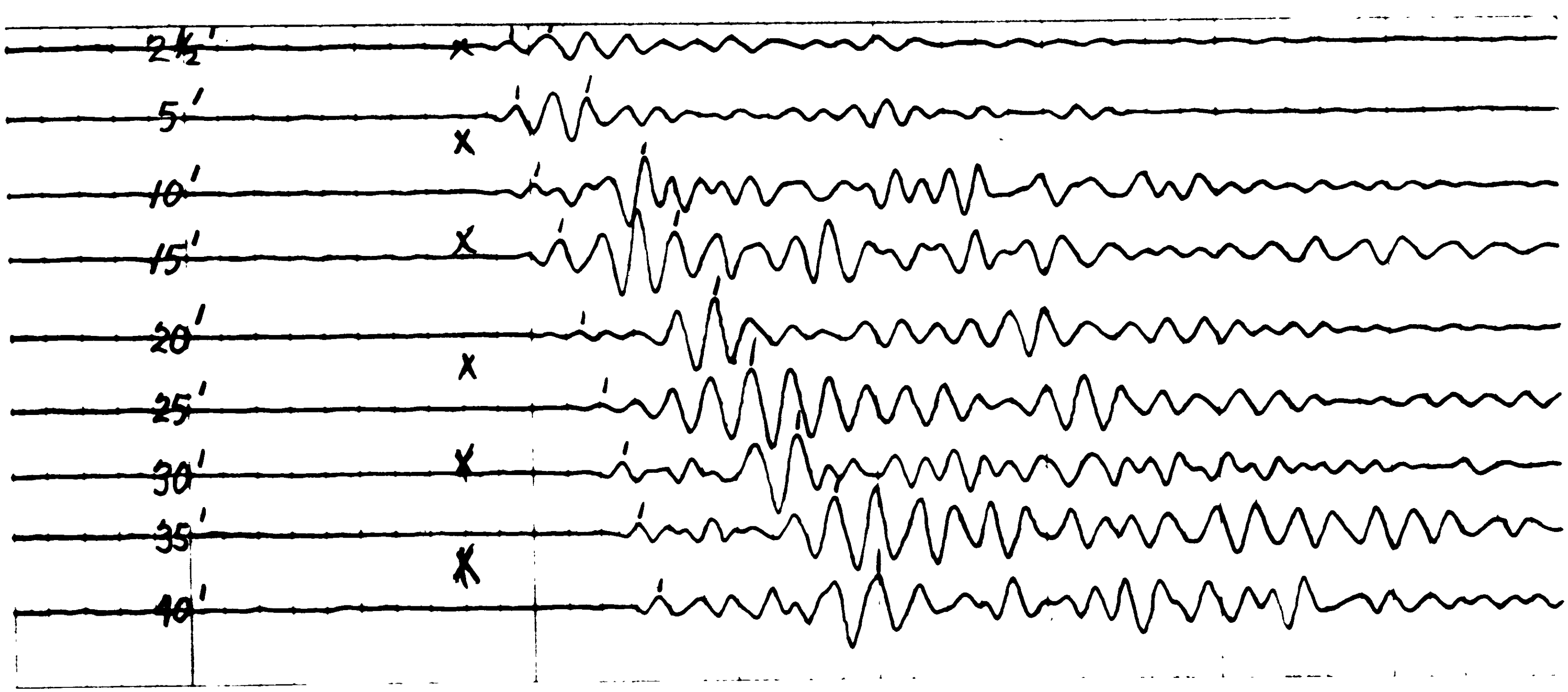


Fig. 3.3

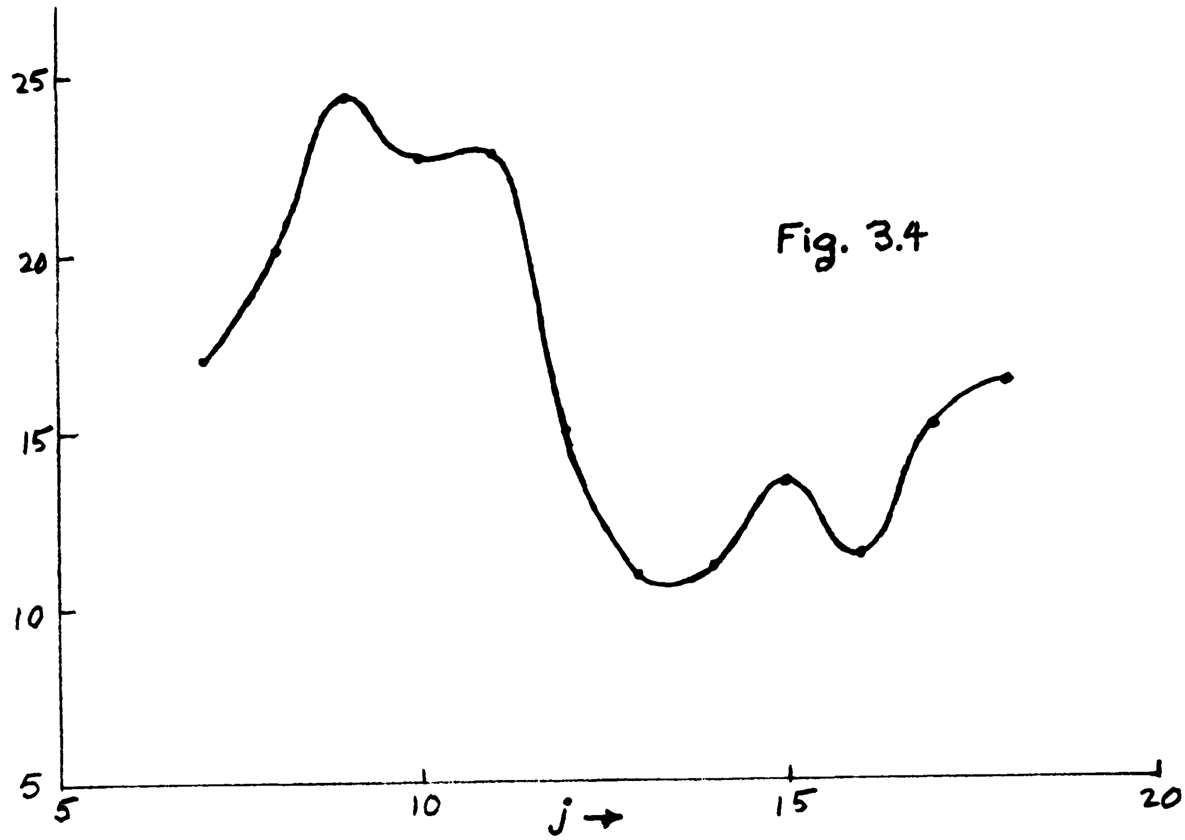


Fig. 3.4

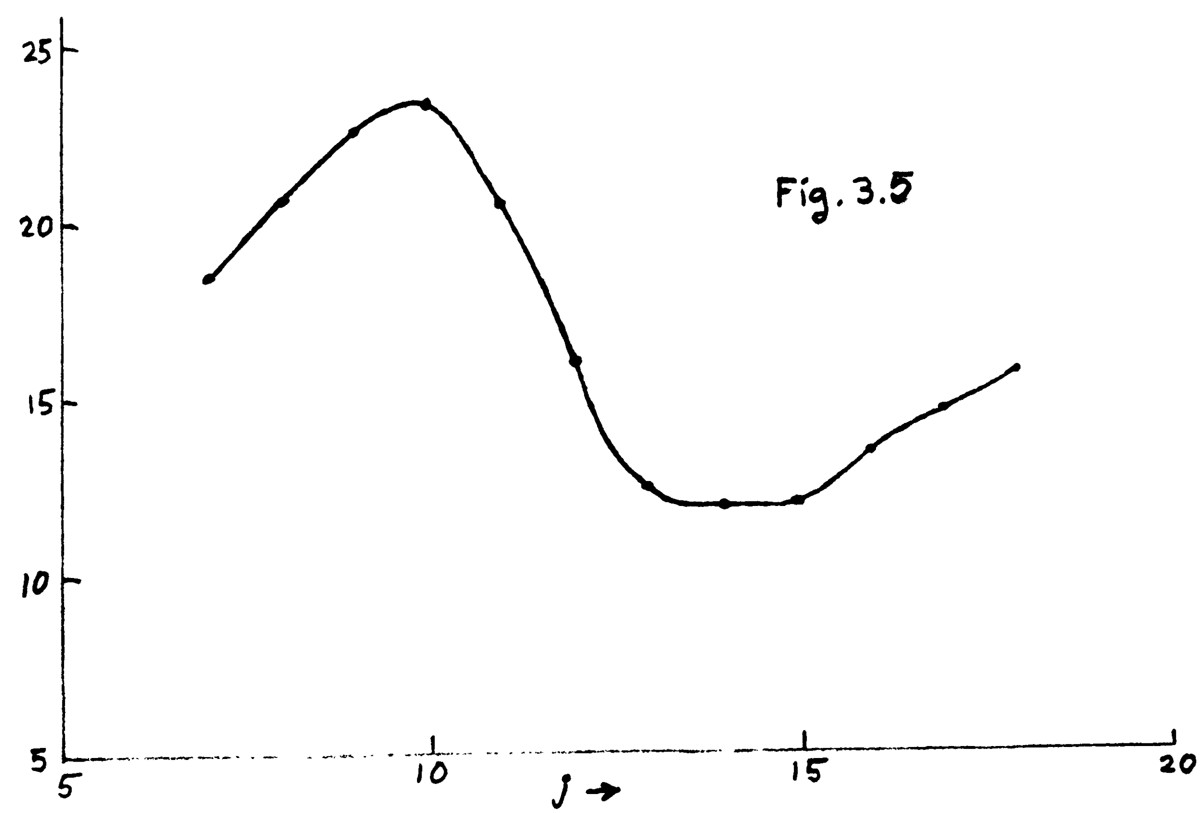


Fig. 3.5

corresponding velocity was quite close to that computed from the second line of check marks.

2. Predicting a Reflected Wave

To test whether or not a reflection could be predicted by these methods, a seismogram showing a prominent reflection was chosen (MIT Record No. 1 ‡). In this record linear operators had been computed and error curves derived. These error curves showed marked peaks at the reflection so the curves were taken as a basis of comparison.

From two traces on this record equation 3.5 was computed. j was selected from the initial step-out between the traces and the non-reflection interval chosen to occur after the reflection. The variance of the predicted wave (in overlapping groups of ten) is plotted in Fig. 3.6 . The dotted and dashed curves of this figure show error curves for linear operators with different prediction distances k. The variance curve does not reach a peak in the reflection as rapidly as do the error curves, but it does compare favorably with them in general shape during and after the reflection. Before the reflection the discrepancy is more noticeable. This may very well be attributable to the fact that operator interval was chosen just before the reflection. In the operator interval, the least squares fitting procedure forces the error curves to be as low as possible.

‡ Ref. 4

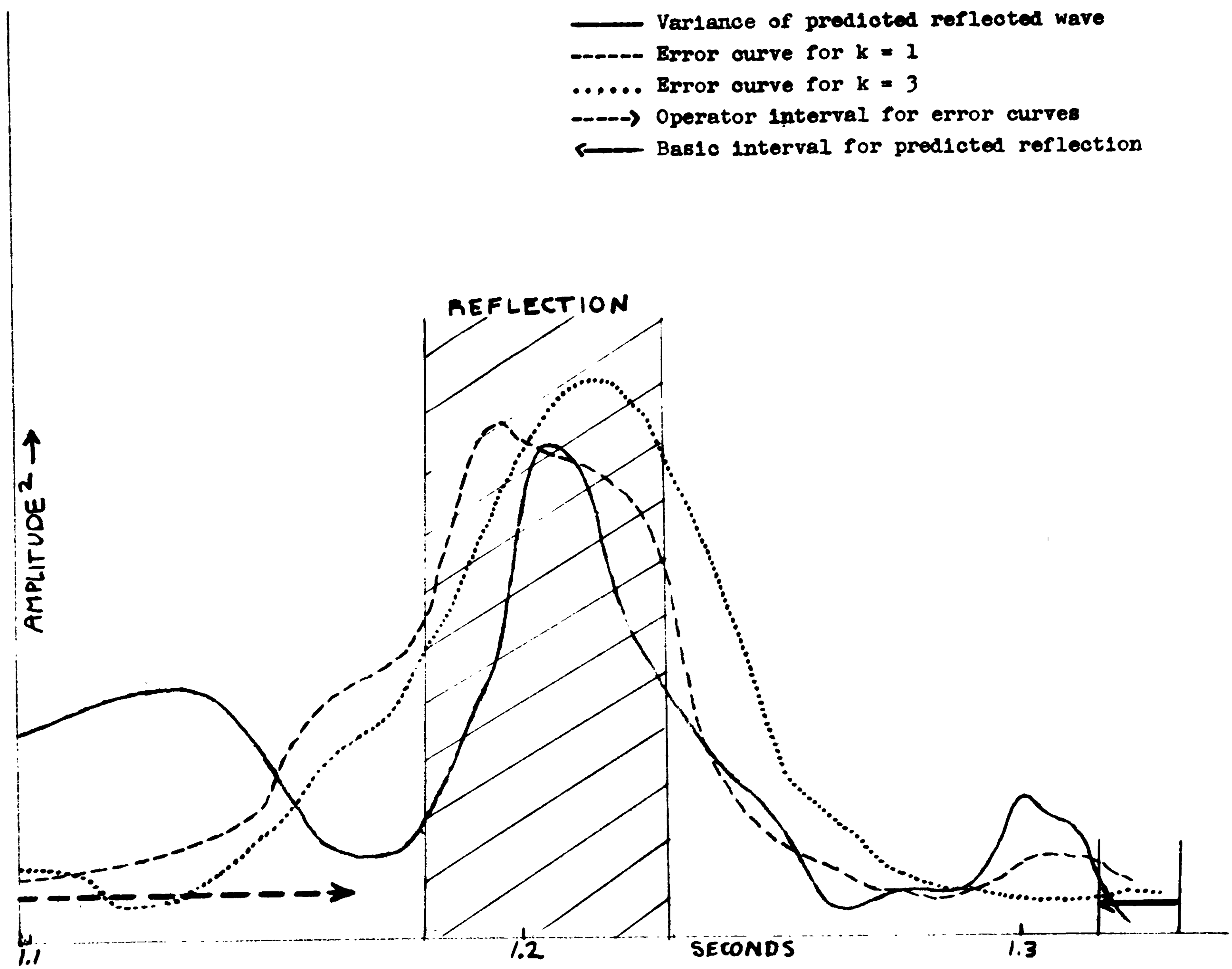

Variance of predicted reflected wave
Error curve for k = 1
Error curve for k = 3
Operator interval for error curves
Basic interval for predicted reflection
REFLECTION
AMPLITUDE²
1.1
1.2
SECONDS
1.3

Fig. 3.6

Conclusions

The two tests discussed show that the expected effects are noted. However, the data used are reasonably ideal, in the sense that ordinary methods of interpretation are adequate. Whether or not the statistical techniques are better can only be determined by many further trials. Situations difficult to treat by the ordinary methods will also fail to uphold the simple assumptions of the theory presented here. On the other hand, only the simplest forms of the theory were used in the examples. Refinements, such as the use of three or more traces for reflection picking may give more valid results.

Phase Test

The "error curve", as used by the Geophysical Analysis Group for picking reflections, is a running average of the squared differences between a predicted and an actual seismic trace. Fig. 3.7 shows an actual trace (the solid curves), and three predictions of this trace, from linear operators with different values of prediction distance. From this diagram we see that the error curve is a running measure of the vertical differences between the predicted and actual traces.

At the reflection (shaded) these differences are seen to become large, and hence the error curve rises to a peak in this interval. The reason the differences become large is not because there is a big discrepancy between the average amptitudes of the predicted and actual traces. From the diagram it appears that the reason is that there is a horizontal displacement of the oscillations of one trace with respect to the other. In other words, there is a phase shift between the predicted and actual traces during the reflection, which disappears shortly after the reflection.

It seems then that a test of phase relationships might well show the reflections as well as the error curve does. A fairly rigorous way of testing this phase shift would be the following:

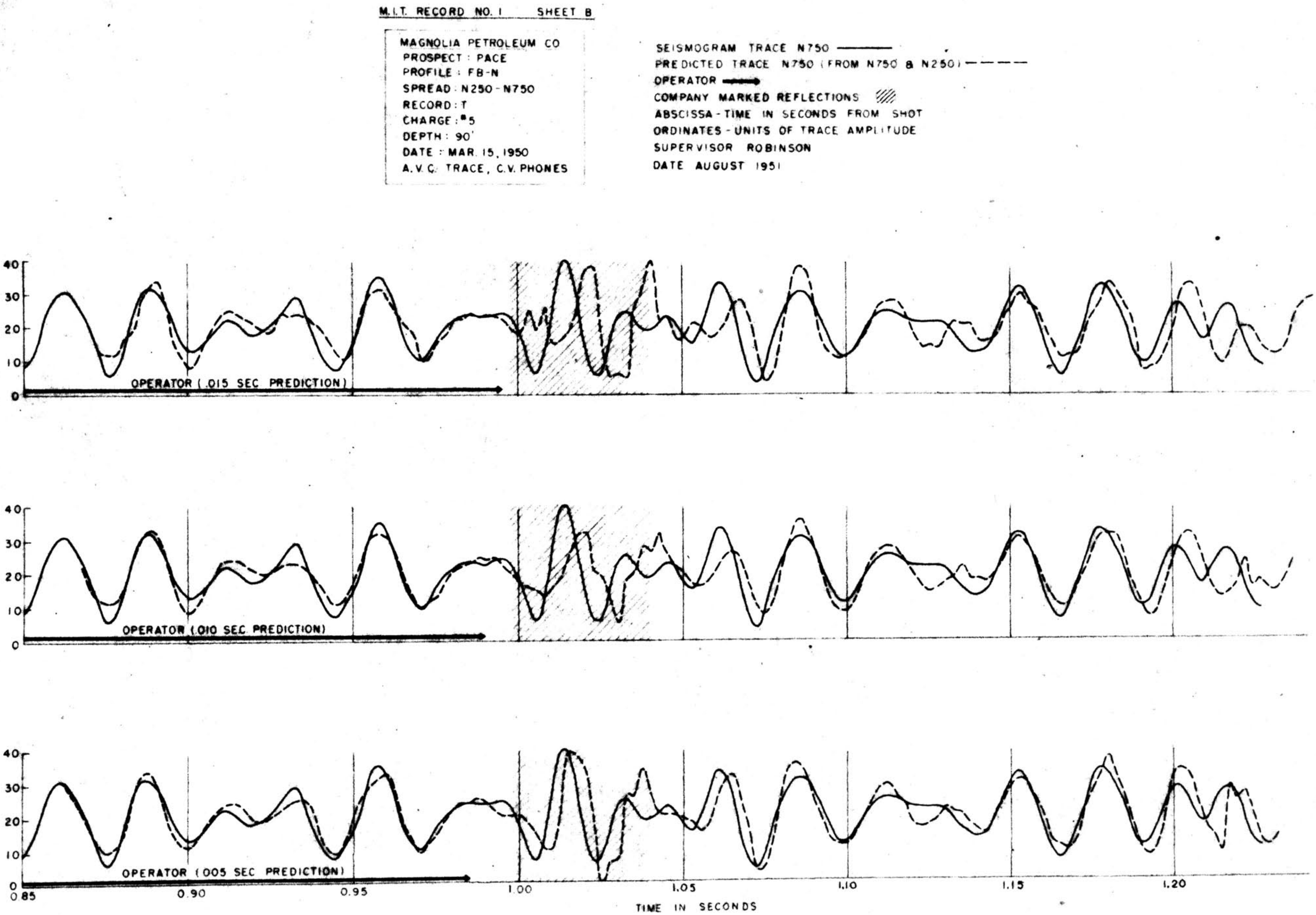

M.I.T. RECORD NO. I SHEET B
MAGNOLIA PETROLEUM CO
PROSPECT: PACE
PROFILE: FB-N
SPREAD: N250-N750
RECORD: T
CHARGE: #5
DEPTH: 90'
DATE: MAR. 15, 1950
A.V.C. TRACE, C.V. PHONES
SEISMOGRAM TRACE N750
PREDICTED TRACE N750 (FROM N750 & N250)
OPERATOR
COMPANY MARKED REFLECTIONS
ABSCISSA - TIME IN SECONDS FROM SHOT
ORDINATES - UNITS OF TRACE AMPLITUDE
SUPERVISOR ROBINSON
DATE AUGUST 1951
OPERATOR (.015 SEC PREDICTION)
OPERATOR (.010 SEC PREDICTION)
OPERATOR (.005 SEC PREDICTION)
0.85
0.90
0.95
1.00
1.05
1.10
1.15
1.20
TIME IN SECONDS

Fig. 3.7

1. Select highly overlapping intervals of the record.

2. Compute the cross-spectrum between the predicted and actual traces in each interval, thus obtaining the phase relationships.

3. Plot the phase angle of the dominant frequency as a function of the interval chosen.

Practically, this is an involved procedure. We can use a simple but crude method to get approximately the same results. Since phase shift is expressed by horizontal displacement we can measure this displacement directly from graphs such as in Fig. 3.7 . This requires that we be able to follow corresponding waves in the two traces, which is subject to personal interpretation.

The displacement was measured for the upper set of curves in Fig. 3.7 . From equally spaced points (in time) on the solid curve the horizontal displacements to the dashed curve were measured. Displacements to the right were considered positive, those to the left negative. Where such measurements could not logically be made (for example on the peak occurring at about .96 sec.) values were taken midway between the last value that could logically be made and the next such value. Once this series of displacements was determined, its individual terms were summed in groups of twenty overlapping by ten, in order to smooth the data. These sums are plotted in Fig. 3.8 .

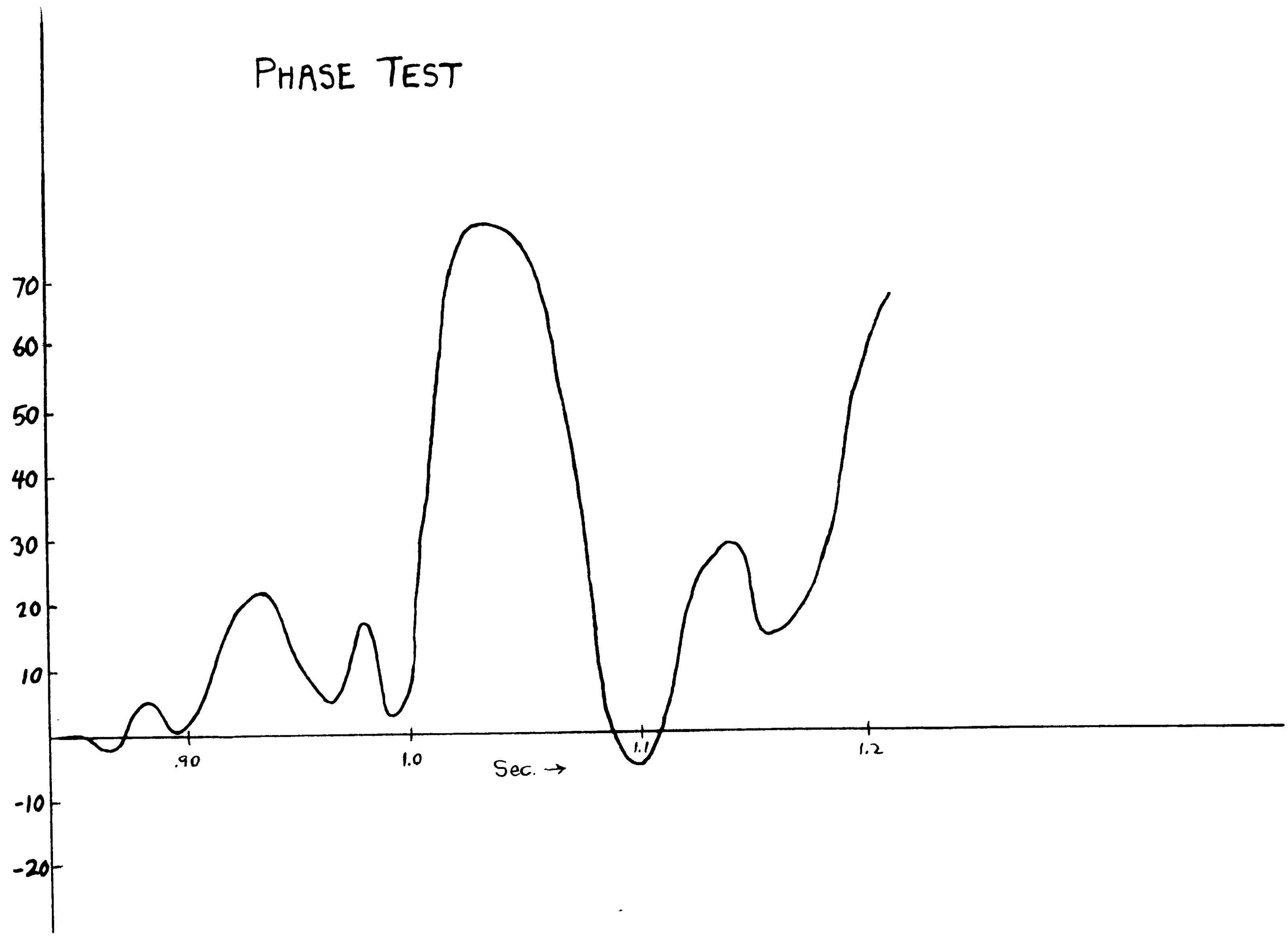


Fig. 3.8

This curve indicates a rapid change of phase occurs at the reflection, the phase rising to a peak in the middle of the reflection, and falling off more gradually thereafter. It seems surprising that the curve is almost entirely positive. If this effect is characteristic, perhaps we should consider as significant only those portions of the curve above a certain mean (about 25 or 30 units in Fig. 3.7). From the original record it appears that there may be another reflection at about 1.23 seconds, which could conceivably cause the rise at the end of the curve.

This is a purely empirical curve. Perhaps it only holds for the particular case treated. One would suspect that the arrival of reflected energy would be accompanied by a rapid change in phase relationships. However, it does not seem reasonable that these changes should be in one direction since the times of arrival of reflected energy are random. Possibly we should deal with the original series only, and compute the <u>rate of change</u> of phase angle (between two overlapping intervals) as a function of interval.

Ensemble Average

The step-out time of a reflection is a property of several seismic traces rather than just a single trace. The error curve for linear operators, as defined elsewhere in this paper, is a property of a single trace - a time average of a single error time series. To get information on the step-out time we must consider operators chosen for different traces. In this connection it is convenient to use an "ensemble" average. This is an average across the "ensemble" of error time series generated by the various operators chosen.

Let us suppose that we have taken a series of operators on a record which consists of traces from equally spaced seismometers. Suppose there are T traces, and on the ℓ^{th} trace ($\ell = 1, 2, \ldots T$) we have chosen N_ℓ operators. For the k^{th} operator on this trace ($k = 1, \ldots N_\ell$) there is an associated error time series which we define as $e_1^{(k\ell)}$. Then, for example, we may construct a single error time series ϵ_1^ℓ to be associated with the ℓ^{th} trace by the expression

$$\epsilon_1^\ell = \sum_{k=1}^{N_\ell} [e_1^{(k\ell)}]^2 \qquad 3.7$$

We may then average these error time series over the various traces. Between traces we observe the effect of step-out. Hence we construct the error time series $\delta_1^{(\alpha)}$ with an arbitrary lag or lead α

$$\delta_{i}^{(\alpha)} = \sum_{\ell=1}^{T} \epsilon_{i-\alpha\ell}^{(\ell)} \qquad \alpha = 0, \pm 1, \pm 2, \ldots \quad 3.8$$

with the expectation that a peak on this error time series, corresponding to a certain reflection, should be highest and narrowest for that value of α most closely corresponding to the true step-out of the given reflection.

No attempt has been made yet to compute 3.8 . It would be a fairly simple task to program this equation for the WWI Digital Computer as a follow-up of the Prediction XV program described in Appendix B.

Travelling Correlations

As an approach to the problem of using a special geophone layout for reflection picking, consider the following arrangement. Two geophones G_1 and G_2 are placed in the ground, one vertically under the other at a distance d. Assuming the ground homogeneous and non-dispersive around the geophones, the responses of G_1 and G_2 may be considered to be due to superpositions of many plane waves travelling with a velocity V from many different directions. In the absence of big reflections, the major contribution to the responses will come from waves having directions not far from the horizontal.

Now consider the cross-correlation of the two responses at G_1 and G_2. In particular consider the value of the function for a time lag equal to d/V. It appears that this value will be strongly influenced by the amount of vertical wave contribution present in the responses, since d/V is the time of direct travel from G_2 to G_1. The cross-correlation at the lag d/V should rise rapidly at a reflection and drop off afterward.

In practice we would have to compute this correlation over highly overlapping time intervals of the response functions in order to obtain the correlation as a function of time. The correlation program described in Appendix D is adaptable to this type of analysis. So far however, no seismograms with the above geophone arrangement have been available.

CONCLUSIONS AND SUGGESTIONS FOR FUTURE WORK

It is difficult to make evaluations of validity on methods which have undergone little testing. Nevertheless we may draw certain conclusions from the work presented here.

Polynomial gravity approximation, as presented here, seems of sufficient simplicity and validity to justify a considerable amount of further study. If further trials show more promise it would be well worth-while to find the inverses of the matrices of Table III. In any event, polynomial approximations of this type have applications in many other fields, and the simplifications brought forward here may be of real value in these other applications.

The properties of cosine operators are of mathematical interest, but it is hoped that studies of this sort will lead to more practical results. In particular, further pursuit of the filter characteristics of linear operators will lead to a better understanding of the extent of realizability of equivalent electronic filters, and or to simplification in the determination of such operators.

The author is more hesitant about recommending the various procedures discussed in Part III. Seismograms exhibit extreme variability in their characteristics and, whereas the examples given here are encouraging, the procedures may fail on other types of records. However, the problems they attempt to settle are of great practical concern and all promising techniques should be either proved or disproved. Phase is a crucial variable in these problems, and probably considerable effort should be spent studying this parameter.

As for the Appendixes, the author feels that the programs described therein have genuine value. Anyone concerned with research depending largely on computation appreciates the fact that obtaining errorless results is a major problem. Programs such as these effectively eliminate this type of problem, and are available for the use of persons interested in the sort of computations they perform.

REFERENCES

1. Agocs, W.B., "Least Squares Residual Anomaly Determination," Geophysics XVI, 4, October, 1951

2. Crout, P:D., "A Short Method for Evaluating Determinants and Solving Systems of Linear Equations with Real or Complex Coefficients," Trans. Am. Inst. of Elect. Eng., Vol. 60, 1941

3. Tukey, J:W., "The Sampling Theory of Power Spectrum Estimates," Woods Hole Symp. Autocorr. Applic. to Phys. Prob., pp. 46 - 67 D.N.R. NAVEXOS - P - 735, 1949

4. Wadsworth, G.P., and Robinson, E.A., Directors, M.I.T. G.A.G. Report No. 1, "Results of an Auto-Correlation and Cross-Correlation Analysis of Seismic Records, Presented at M.I.T. Industrial Liason Conference of August 6, 1952," October 1, 1952

5. Wadsworth, G.P., and Robinson, E.A., Directors, M.I.T. G.A.G. Report No. 2, "A Prospectus on the Application of Linear Operators to Seismology," November, 1952

6. Wadsworth, Robinson, Hurley, "Detection of Reflections on Seismic Records," Geophysics XVIII, 3, July, 1953

7. Kendall, Maurice G., The Advanced Theory of Statistics Vol. II, Charles Griffin and Compnay Limited, London, 1947

8. Southwick, Peter, Velocity of Shear Waves Through Unconsolidated Materials, M.I.T. Ph.D. Thesis, 1952

APPENDIX A

APPENDIX A

POLYNOMIAL I (2492 m 1) - Description and Use Of

This program was written for the WW1 Digital Computer to eliminate the task of computing the residuals from a least squares fitting of an n^{th} order polynomial to data taken over an arbitrary-sized rectangle. A copy of the program appears at the end of this Appendix.

Polynomial I does the following:

I. It solves the equation

$$
\begin{aligned}
g(xy) = {} & C_{00} + C_{10}x + C_{20}x^2 + \ldots\ldots + C_{n0}x^n \\
& + C_{01}y + C_{11}xy + \ldots\ldots + C_{(n-1),1}x^{n-1}y \\
& \quad \vdots \qquad\quad \vdots \\
& + C_{0,(n-1)}y^{n-1} + C_{1,(n-1)}xy^{n-1} \\
& + C_{0n}y^n
\end{aligned}
$$

where the $C_{\ell k}$'s are given, n, the order of the polynomial is given, and the values of x and y are to be taken over a rectangular grid measuring 2N by 2M, and with axes centered as in Fig. 1.2.

II. It then forms the differences g(xy) for all values of g(xy) on the grid. These are the residuals.

III. It prints out these residuals in the same network fashion that the grid was chosen.

Use of Polynomial I

There are certain conventions which must be observed in the use of this program. The constants defining the nature of the polynomial and grid must appear as follows:

Register	440	+n	order of polynomial (less than 7)
(Octal)	441	+N	greatest value of x
	442	+M	greatest value of y

The coefficients $C_{\ell k}$ of the polynomial must be scale factored in a special way because they decrease in magnitude rapidly as $\ell+k$ increases. The scale factor is $10^{(\ell+k)-2}$, which in most instances will guarantee that all are less than unity in absolute value, but not greatly so. They must appear in the machine as follows:

Register	443	C_{00}	x	10^{-2}	461	C_{12}	x	10
(Octal)	444	C_{10}	x	10^{-2}	462	C_{22}	x	10^{2}
	445	C_{20}	x	1	463	C_{32}	x	10^{3}
	446	C_{30}	x	10	464	C_{42}	x	10^{4}
	447	C_{40}	x	10^{2}	465	C_{03}	x	10
	450	C_{50}	x	10^{3}	466	C_{13}	x	10^{2}
	451	C_{60}	x	10^{4}	467	C_{23}	x	10^{3}
	452	C_{01}	x	10^{-1}	470	C_{33}	x	10^{4}
	453	C_{11}	x	1	471	C_{04}	x	10^{2}
	454	C_{21}	x	10	472	C_{14}	x	10^{3}

455	C_{31}	x	10^2	473	C_{24}	x	10^4
456	C_{41}	x	10^3	474	C_{05}	x	10^3
457	C_{51}	x	10^4	475	C_{15}	x	10^4
460	C_{02}	x	1	476	C_{06}	x	10^4

The data g(xy) which is presumed to be taken over the gridwork, is scale factored by 10^{-2} and appears in the machine as follows:

Register	540	g(-N,M)	x	10^{-2}
	541	g(-N+1,M)	x	10^{-2}
	542	g(-N+2,M)	x	10^{-2}
	.	.		
	.	.		
	.	g(N,M)	x	10^{-2}
	.	g(-N,M-1)	x	10^{-2}
	.	g(-N+1,M-1)	x	10^{-2}
	.	.		
	.	.		
	.	g(N,M-1)	x	10^{-2}
	.	g(-N,M-2)	x	10^{-2}
	.	.		
	.	.		
	.	.		
	.	g(-N,-M)	x	10^{-2}
	.	g(-N+1,-M)	x	10^{-2}
	.	.		
	.	.		

. $g(N,-M)$ x 10^{-2}

. $g(-N+1,-M)$ x 10^{-2}

.   .

.   .

. $g(N,-M)$ x 10^{-2}

Now suppose the information n, N, M, and the $C_{\ell k}$'s are prepared on a tape with the tape number X, and the date g(xy) is prepared on a tape with the tape number Y. Then the instructions for the operation of this program would be

Erase storage

Read in 2492 m 1

Read in X

Read in Y

Start at 127 (Octal)

The residuals are printed out by the direct printer in about three minutes or so depending on the size of the grid. They appear as four-digit numbers where the decimal point is understood to occur after the second digit.

As an example of the output we include a sample of three sets of residuals. These were derived for the three sets of coefficients used elsewhere in this paper. The sample illustrates the convenience of this form of answer for contouring purposes. In fact, with only slight modification (inserting two extra carriage returns between lines) these numbers would appear on a grid with square unit cell, and could be contoured directly, on the result sheet.

<u>A Technical Feature in Polynomial I</u>

We describe here a technical feature in this program which might be of use to other programmers. The problem is that we are multiplying numbers rapidly decreasing in magnitude with $\ell+k$ (the $C_{\ell k}$'s) by numbers rapidly increasing in magnitude with $\ell+k$ ($x^{\ell}y^{k}$) while the product is of a relatively constant order of magnitude, which must be in a form which we can add to other such products.

What we want is the product $C_{\ell k}x^{\ell}y^{k}$ to be scale factored finally by 10^{-2}. To preserve accuracy during the computation of the product we do the following:

1. Form $x^{\ell} \quad 2^{-15}$ and $y^{k} \quad 2^{-15}$ and then scale factor to $x^{\ell} \quad 2^{-15+\alpha}$ and $y^{k} \quad 2^{-15+\beta}$ by use of the scale factor order.
2. Form $C_{\ell k}10^{(\ell+k)-2}x^{\ell}2^{-15+\alpha}y^{k}2^{-15+\beta}$

$$= C_{\ell k}x^{\ell}y^{k}2^{-30+\alpha+\beta}10^{(\ell+k)-2} \qquad (1)$$

To get this product to the form $C_{\ell k}x^{\ell}y^{k}10^{-2}$ we must multiply by $10^{-(\ell+k)}2^{30-(\alpha+\beta)}$

It appears that we merely need to store the negative powers of 10, multiply the expression (1) by $10^{-(\ell+k)}$ and then shift left $30-(\alpha+\beta)$. However the negative powers of 10 cannot be stored with any accuracy for high $\ell+k$ so we write $10^{-(\ell+k)}2^{30-(\alpha+\beta)}$ in the form

$$2^{-(\ell+k)\log_2 10+30-(\alpha+\beta)}$$

$$= 2^{-3.32193(\ell+k)+30-(\alpha+\beta)}$$

$$= \frac{1}{2^{.32193(\ell+k)}} \quad [2^{-3(\ell+k)+30-(\alpha+\beta)}]$$

$$= \frac{1}{(2^{.32193})^{(\ell+k)}} [2^{-3(\ell+k)+30-(\alpha+\beta)}]$$

$$= (.800)^{\ell+k} [2^{-3(\ell+k)+30-(\alpha+\beta)}]$$

We can store the powers of (.800) with ample accuracy. Thus we multiply by the appropriate power of (.800) and follow this by a shift left or right according to the exponent of 2. (The zeroth power of (.800) is put in as +.9999.)

m = 4

+0385	-0302	-0008	-0147	-0094	-0136	+0082	+0153	+0209	+0173	-0146
+0163	+0004	-0325	+0062	-0049	-0193	-0007	-0045	+0031	+0001	-0120
+0086	-0132	-0082	+0183	+0553	+0301	-0056	-0116	+0063	+0069	-0078
+0039	-0083	-0097	+0250	+0199	-0115	-0176	-0122	-0003	+0017	-0030
-0180	-0195	+0032	+0311	-0058	-0043	-0014	-0158	-0018	+0049	-0021
-0203	-0151	+0025	+0141	+0050	-0007	-0247	-0250	+0145	+0141	+0074
-0145	-0164	-0079	+0376	+0563	+0179	-0186	-0208	-0220	+0084	+0149
+0099	-0028	+0070	+0071	+0138	+0175	-0222	-0124	-0158	+0087	+0165
+0249	+0225	+0098	-0149	-0151	-0045	-0231	-0123	-0041	+0030	+0098
+0095	-0060	-0004	-0097	-0264	-0141	-0070	-0010	+0173	+0104	+0044
-0157	+0417	-0482	-0112	+0106	+0097	+0220	+0224	+0350	+0125	-0583

m = 3

+0372	-0293	-0001	-0172	-0181	-0300	-0154	-0115	-0007	+0146	+0223
+0193	+0085	-0211	+0179	+0035	-0166	-0051	-0140	-0058	+0021	+0175
+0067	-0087	+0020	+0318	+0684	+0393	-0027	-0150	-0000	+0050	+0078
-0057	-0114	-0057	+0343	+0313	-0021	-0132	-0145	-0081	-0065	-0021
-0333	-0299	+0003	+0350	+0019	+0036	+0027	-0179	-0108	-0082	-0124
-0363	-0292	-0050	+0138	+0100	+0063	-0196	-0253	+0070	+0001	-0084
-0249	-0296	-0170	+0355	+0605	+0259	-0108	-0169	-0248	-0018	-0002
+0107	-0112	-0008	+0047	+0181	+0270	-0108	-0030	-0117	+0056	+0070
+0404	+0205	+0033	-0184	-0122	+0046	-0100	+0012	+0062	+0075	+0079
+0390	-0041	-0095	-0193	-0309	-0115	+0014	+0106	+0288	+0186	+0076
+0212	+0390	-0700	-0382	-0131	-0067	+0133	+0200	+0358	+0134	-0595

m = 2

+0105	-0402	-0035	-0189	-0217	-0368	-0242	-0189	-0009	+0297	+0632
+0044	+0079	-0158	+0231	+0048	-0208	-0138	-0240	-0115	+0085	+0464
-0000	-0019	+0136	+0421	+0731	+0369	-0115	-0273	-0104	+0041	+0264
-0075	-0001	+0099	+0478	+0383	-0033	-0224	-0288	-0224	-0134	+0082
-0333	-0164	+0181	+0504	+0104	+0030	-0065	-0334	-0277	-0192	-0080
-0372	-0158	+0130	+0298	+0192	+0063	-0288	-0413	-0111	-0131	-0075
-0293	-0185	-0001	+0510	+0699	+0264	-0192	-0323	-0426	-0152	-0002
+0003	-0043	+0134	+0190	+0273	+0283	-0179	-0165	-0274	-0056	+0088
+0218	+0215	+0138	-0061	-0033	+0070	-0148	-0089	-0054	+0007	+0147
+0102	-0106	-0039	-0093	-0222	-0073	+0002	+0054	+0234	+0191	+0225
-0195	+0239	-0698	-0309	-0042	-0000	+0169	+0217	+0391	+0243	-0328

	Address	Contents	Remarks
	00		
	01		
	02		
	03		
	04		
	05		
	06		
	07		
	10		
	11		
	12		
	13		
	14		
	15		
	16		
	17		
	20		
	21		
	22		
	23		
	24		
	25		
	26		
	27		
	30		
	31		
	32		
	33		
	34		
	35		
	36		
	37		
	40	ta 63	↑
	41	ta 70	
	42	si 225	
	43	ts 113	
	44	sp 114	PRINT
	45	ts 110	
	46	cm 113	
	47	mh 71	
	50	ad 72	
	51	td 54	
	52	sr 17	ROUTINE
	53	ts 113	
	54	ca	
	55	rc 55	
	56	ao 110	
	57	cp 46	
	60	ca 105	
	61	rc 61	
	62	ao 112	
	63	cp	
	64	ca 106	
	65	rc 65	
	66	ca 111	
	67	ts 112	
	70	sp	
	71	+10	
	72	0.00073	
	73	1.74000	
	74	0.52000	
	75	0.36000	
	76	0.16000	
	77	0.26000	

	Address	Contents	Remarks
	100	0.46000	
	101	0.66000	
	102	0.56000	
	103	0.06000	
	104	1.54000	
	105	0.20000	
	106	1.22000	
	107	+3	
	110	si 0	
	111	si 0	
	112	si 0	
	113	si 0	
	114	cp 123	
	115	ca 121	
	116	rc 116	
	117	ct 107	
	120	sp 45	
	121	0.32000	
	122	0.72000	
	123	ca 122	
	124	sp 116	
	125	si 0	
	126	si 0	↓
START →	127	ca 441	↑
	130	ts 322	
	131	su 441	
	132	ts 317	
	133	ts 112	SET
	134	ca 442	
	135	su 442	
	136	ts 321	
	137	ca 442	REGISTERS
	140	ts 325	
	141	ca 341	
	142	td 300	
	143	ao 325	
	144	ts 326	
	145	ca 317	
	146	ts 320	
	147	ca 326	
	150	su 312	
	151	ts 326	COUNTERS
	152	ca 322	
	153	su 312	
	154	ts 323	
	155	ao 323	
	156	ca 312	
	157	ts 324	
	160	ts 327	
	161	ca 331	
	162	td 230	ADDRESSES
	163	ca 440	
	164	ts 313	
	165	ts 314	
	166	ts 315	
	167	ts 316	
	170	ca 340	
	171	td 266	
	172	ca 364	
	173	td 233	
	174	ca 0	
	175	ts 343	
	176	ts 345	
	177	ca 376	

	Address	Contents	Remarks
	200	td 253	↓
	201	ca 324	↑
	202	sf 346	
	203	ts 350	
	204	ca 327	FORM
	205	sf 347	
	206	ts 351	
	207	ca 343	
	210	su 345	
	211	mh 353	ĝ(n,m)
	212	slr 17	
	213	ad 352	MINUS
	214	su 346	
	215	su 347	
	216	cp 223	
	217	td 373	
	220	ca 373	
	221	ts 234	g(n,m)
	222	sp 230	
	223	mh 375	
	224	slr 17	
	225	td 374	
	226	ca 374	
	227	ts 234	
	230	ca 0	
	231	mh 351	
	232	mh 350	
	233	mh	
	234	si 0	
	235	ad 330	
	236	ts 330	
	237	ca 324	
	240	mh 323	
	241	slr 17	
	242	ts 324	
	243	ao 230	
	244	ao 233	
	245	ao 343	
	246	ao 314	
	247	cp 201	
	250	ao 345	
	251	ca 0	
	252	ts 343	
	253	ca 0	
	254	sp 256	
	255	ca 0	
	256	td 233	
	257	ao 253	
	260	ca 327	
	261	mh 326	
	262	slr 17	
	263	ts 327	
	264	ca 312	
	265	ts 324	
	266	ca 0	
	267	sp 271	
	270	ca 0	
	271	td 230	
	272	ao 266	
	273	ao 313	
	274	ts 314	
	275	ao 316	
	276	cp 201	
	277	ca 330	

	Address	Contents	Remarks
	300	ad 0	
	301	sp 40	
	302	ca 0	
	303	ts 330	
	304	ao 300	
	305	ao 320	
	306	cp 155	
	307	ao 321	
	310	cp 145	
	311	si 0	PROGRAM END ↓
	312	+1	↑
	313	si 0	
	314	si 0	
	315	si 0	
	316	si 0	
	317	si 0	ADDRESSES
	320	si 0	
	321	si 0	
	322	si 0	
	323	si 0	
	324	si 0	
	325	si 0	
	326	si 0	
	327	si 0	
	330	si 0	
	331	ca 443	COUNTERS
	332	ca 452	
	333	ca 460	
	334	ca 465	
	335	ca 471	
	336	ca 474	
	337	ca 476	
	340	ca 332	
	341	ca 540	
	342	si 0	CONSTANTS
	343	si 0	
	344	si 0	
	345	si 0	
	346	si 0	
	347	si 0	
	350	si 0	
	351	si 0	
	352	+30	
	353	+3	
	354	si 0	
	355	+.9999	
	356	+.8000	
	357	+.6400	
	360	+.5120	
	361	+.4096	
	362	+.3276	
	363	+.2621	
	364	ca 355	
	365	ca 356	
	366	ca 357	
	367	ca 360	
	370	ca 361	
	371	ca 362	
	372	ca 363	
	373	slr 0	
	374	srr 0	
	375	-1	
	376	ca 365	↓
	377		

NOTES

REG 440 +n	452 C_{01}		REG 540 $g(-N, M) \times 10^{2}$
441 +N	453 C_{11}	476 C_{06}	541 $g(-N+1, M)$
442 +M	:	C_{lk}	:
443 C_{00}	460 C_{02}	are $\times 10^{(l+k)-2}$	$g(-N, M-1)$
444 C_{10}			$g(-N+1, M-1)$

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE POLYNOMIAL I INDEX
AUTHOR SIMPSON DATE
TAPE NUMBER 2492-m1

DL-481

APPENDIX B

APPENDIX B

PREDICTION XV (2539 m 2) - Description and Use Of

This program was written to provide computational facility for predicting a series $x_i(y_i, z_i,$ or $u_i)$ with a linear operator of the general form

$$x_{i+k}(y_{i+k},\ z_{i+k} \text{ or } u_{i+k}) = a + a_0x_i + a_1x_{i-1} \ldots + a_Mx_{i-M}$$
$$+ b_0y_i \ldots\ldots\ldots\ldots + b_My_{i-M}$$
$$+ c_0z_i \ldots\ldots\ldots\ldots + c_Mz_{i-M}$$
$$+ d_0u_i \ldots\ldots\ldots\ldots + d_Mu_{i-M}$$

where the prediction distance k and the number of lags M are arbitrary but have the restrictions that M 7 and M+k 19. The four series which this program handles contain 500 members each so that i ranges from o through 499. This first prediction computed is for i+k = 20 and the last one for i+k = 499. After doing this computation the program forms the running average of squared errors between the predicted and actual series

$$\sum_{i=j-5}^{j+4} (x_i - x_i)^2 \qquad j = 25, 35, 45, 55, \ldots 495$$

which is called the "Error Curve".

There is considerable choice of output. The alternatives are any combination of or none of the following:

1. Print-out of the errors and sums of squares
2. Print-out of just the sums of squares of errors.

3. Photographs of oscilloscope displays of the sums of errors squares

An additional choice is the use of magnetic tape delayed output for 1 and 2 above, which is about fifteen times faster than direct print-out.

This program handles up to eight operators at a time in the above fashion. When the magnetic tape output is used, the error curves can be removed from the machine at the rate of one every ten seconds whereas the individual errors and error curves require fifty seconds for each operator. Each curve would represent about a week of hand computation. Once the computations are completed the individual errors (x_i-x_i) for all operators are left in magnetic drum storage so other programs can use them for different types of averaging processes rather than just the Error Curve as described above.

On the next three pages are illustrated the various output forms. The first page is a reproduction of the individual errors and error curves for two operators. The results for each operator appear as a block of numbers 10 by 48 and a right-hand column of 48 numbers. The block represents the 480 individual errors whereas the right-hand column is the Error Curve, each member representing the sums of the squares of the 10 individual errors in the corresponding row to the left. The number appearing over the upper left corner of each block is a number assigned to the particular operator for identification purposes, and is printed by the program. The number printed over the center of each block was inserted later.

The next page shows the output form for four operators when just the Error Curve is desired. The +0000 identification number indicates that the operator was chosen as the variance operator which has the forn $x_1 = \bar{\bar{x}}$ (means of series). The Error Curve for this type operator becomes the sample variance curve and provides a basis for testing the statistical significance of other operators predicting the x_1 series.

The third page is a photograph taken automatically by the program of an oscilloscope display of one-half of an Error Curve. A vertical and horizontal axis are also displayed.

+1300

12.4 Second Half

+0027	+0052	+0008	-0036	-0009	+0027	+0016	+0024	+0032	+0024	+0007
+0000	+0007	+0001	-0012	-0006	+0014	-0001	-0041	-0000	-0011	+0001
-0020	+0000	-0009	-0014	-0007	-0030	-0031	-0008	+0054	+0073	+0009
+0058	+0053	+0009	-0025	+0032	+0014	+0003	-0016	-0066	-0024	+0012
+0006	+0019	+0062	+0013	-0021	-0015	+0003	-0002	+0034	+0050	+0007
+0010	+0005	+0013	+0011	-0016	-0007	+0008	+0031	+0040	+0055	+0005
+0057	+0030	+0043	+0012	-0067	-0032	-0019	-0028	+0009	+0003	+0011
-0019	+0001	-0008	-0064	-0021	-0043	-0019	+0043	+0020	+0053	+0011
+0088	+0079	+0054	+0041	+0005	-0040	-0046	-0036	-0034	-0049	+0026
-0059	+0043	+0093	-0006	+0034	-0025	-0078	-0009	-0042	-0091	+0031
-0021	-0011	-0005	+0095	+0153	+0056	+0009	-0048	-0022	+0075	+0043
+0000	-0003	+0046	-0001	-0001	+0021	-0003	+0001	+0027	+0048	+0005
+0016	+0008	+0031	+0005	+0014	-0009	-0022	+0001	-0003	+0026	+0001
+0000	+0010	+0054	+0024	+0019	+0025	-0021	-0027	+0013	+0000	+0005
-0005	-0006	+0014	-0016	+0010	+0030	-0014	-0002	-0054	-0041	+0004
-0049	-0073	-0022	-0050	-0042	+0019	+0045	+0042	+0035	+0028	+0017
-0013	+0002	+0018	-0013	-0027	-0018	-0017	+0008	-0019	+0019	+0001
+0047	+0032	+0040	+0020	+0025	+0002	-0048	-0022	-0024	-0028	+0008
+0002	+0025	+0051	+0067	+0079	+0067	+0037	+0035	-0047	-0058	+0025
-0048	-0092	-0036	-0048	-0060	+0013	-0000	+0007	+0019	-0013	+0017
-0029	-0021	-0025	-0010	+0039	+0018	+0033	+0044	-0043	-0051	+0010
-0002	-0035	-0005	+0010	-0022	+0025	+0040	+0054	+0086	+0031	+0014
+0009	-0010	-0060	-0053	-0025	-0049	-0032	+0042	+0046	+0039	+0015
+0017	+0010	-0043	-0031	-0003	-0057	-0039	-0053	-0009	+0070	+0015
+0037	+0017	-0025	-0058	+0001	-0011	-0010	-0012	-0040	-0005	+0006
+0028	+0021	-0001	+0008	+0014	-0014	-0004	+0021	-0002	+0017	+0001
+0050	+0036	+0017	-0007	-0060	-0052	-0041	-0086	-0023	-0010	+0019
+0021	+0060	-0002	+0003	+0000	-0052	-0052	-0041	-0066	-0061	+0018
-0028	-0008	+0013	+0035	+0029	+0034	+0062	+0082	+0067	+0029	+0018
+0005	-0040	-0069	-0059	-0038	-0040	-0025	-0007	-0004	+0025	+0013
+0043	+0028	-0004	-0004	-0014	-0028	-0036	-0027	-0025	-0017	+0005
-0010	-0002	-0002	-0021	+0008	+0026	-0001	+0005	+0022	-0013	+0001
-0021	-0003	-0009	-0011	-0047	-0039	-0051	-0072	-0001	+0010	+0011
+0023	+0036	+0002	-0002	+0014	+0003	-0031	-0010	-0001	-0017	+0002
-0009	+0012	-0234	-0242	+0186	-0140	-0191	-0075	-0075	-0075	+0220
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054
-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	-0075	+0054

+1401

12.4 Second half

-0023	+0016	+0030	+0003	-0000	-0013	-0029	-0013	+0011	-0006	+0001
-0001	-0036	-0052	+0013	+0010	+0033	+0015	+0010	+0007	-0007	+0004
-0021	-0031	-0003	+0000	-0035	-0032	-0055	-0044	-0003	+0075	+0013
+0134	+0130	+0033	-0028	-0064	-0077	-0034	-0014	-0032	-0053	+0050
-0041	+0060	+0069	+0051	+0016	-0047	-0058	-0030	+0008	+0024	+0018
+0022	-0006	-0012	-0004	+0018	+0007	+0005	+0020	+0035	+0073	+0007
+0069	+0035	+0024	-0024	-0081	-0096	-0068	-0046	-0041	-0023	+0030
+0014	+0012	+0004	+0030	+0014	-0024	-0003	+0006	+0004	+0051	+0003
+0071	+0057	+0046	+0007	-0023	-0061	-0083	-0052	-0052	-0063	+0029
-0054	+0058	+0112	+0128	+0120	-0023	-0149	-0158	-0123	-0086	+0119
-0068	-0018	+0029	+0105	+0190	+0209	+0132	+0000	-0096	-0139	+0142
-0084	+0018	+0004	-0015	+0004	-0018	+0021	-0007	+0017	+0079	+0014
+0033	+0001	+0002	-0023	-0012	+0001	+0041	+0015	-0023	+0003	+0003
-0027	-0019	+0019	+0037	+0050	-0011	-0068	-0052	-0023	+0041	+0013
+0065	+0040	+0001	-0043	-0035	-0019	+0008	+0040	-0030	-0073	+0016
-0126	-0126	-0056	-0036	+0025	+0057	+0098	+0100	+0060	+0049	+0065
-0020	-0051	-0027	-0047	-0056	-0049	-0023	-0005	+0011	+0022	+0011
+0014	+0007	+0052	+0047	+0027	+0005	-0000	-0025	-0061	-0017	+0010
+0002	-0007	+0018	+0039	+0057	+0067	+0074	+0029	-0050	-0041	+0019
-0110	-0146	-0082	-0063	-0008	+0029	+0049	+0032	+0003	+0014	+0048
-0037	-0052	-0037	-0029	-0000	+0039	+0078	+0068	+0021	-0020	+0018
-0038	-0050	-0071	-0057	-0003	-0027	-0006	+0085	+0087	+0042	+0028
+0080	-0006	-0083	-0077	-0110	-0095	-0017	+0025	+0061	+0080	+0050
+0086	+0013	-0079	-0083	-0102	-0039	-0016	-0063	-0003	+0018	+0036
+0081	+0135	+0052	+0002	-0090	-0108	-0057	-0061	-0048	-0020	+0055
+0012	+0037	+0034	+0012	+0003	+0020	+0008	-0003	+0006	+0015	+0002
+0010	+0022	+0042	+0018	-0041	-0084	-0120	-0100	-0028	-0002	+0035
+0077	+0090	+0042	+0000	-0060	-0035	-0074	-0075	-0045	-0071	+0037
-0023	+0013	+0046	+0073	+0032	+0021	+0027	+0030	+0055	+0068	+0018
+0007	-0065	-0078	-0088	-0066	-0029	-0030	-0016	-0001	+0019	+0023
+0015	+0012	+0009	-0027	-0023	-0003	-0010	-0011	-0018	-0011	+0001
-0004	+0009	-0010	-0016	+0006	+0014	+0014	+0015	+0000	-0026	+0000
-0043	-0036	-0009	+0001	+0006	+0013	-0045	-0090	-0054	+0007	+0015
+0032	+0017	+0027	+0013	-0040	-0008	-0003	-0025	-0018	-0031	+0004
-0015	+0010	-0219	-0222	-0215	-0213	-0219	-0209	-0201	-0175	+0352
-0255	-0235	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0379
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323
-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	-0180	+0323

+0000

+0454	+0375	+0437	+0366	+0605	+0253	+0518	+0435	+0273	+0648	+0530	+0342
+0339	+0569	+0354	+0465	+0418	+0611	+0357	+0348	+0422	+0288	+0674	+0390
+0396	+0383	+0422	+0384	+0343	+0468	+0429	+0512	+0332	+0355	+0775	+0256
+0483	+0423	+0313	+0593	+0216	+0395	+0399	+0399	+0399	+0399	+0399	+0399

+0000

+0816	+1241	+0509	+1182	+0781	+0725	+0995	+0784	+0681	+0864	+0902	+0783
+0875	+0728	+0834	+0761	+1013	+0594	+1030	+0871	+0437	+1340	+0646	+0771
+1129	+0536	+0977	+0687	+0949	+0520	+0853	+1589	+0155	+1276	+0878	+0532
+1091	+0550	+0848	+1040	+0384	+0900	+0900	+0900	+0900	+0900	+0900	+0900

+0000

+0373	+0337	+0589	+0437	+0484	+0318	+0519	+0374	+0431	+0507	+0351	+0485
+0391	+0394	+0468	+0431	+0386	+0442	+0423	+0337	+0442	+0395	+0539	+0328
+0359	+0406	+0510	+0307	+0372	+0408	+0413	+0437	+0326	+0395	+0427	+0326
+0388	+0506	+0279	+0451	+0424	+0362	+0399	+0399	+0399	+0399	+0399	+0399

+0000

+0800	+0713	+0818	+0886	+0927	+0833	+0680	+0949	+0806	+0843	+0703	+0846
+0691	+0887	+0893	+0804	+0659	+0938	+0780	+0797	+0695	+0823	+1077	+0800
+0653	+0698	+0877	+0663	+0758	+0905	+0747	+0802	+0603	+0802	+0952	+0696
+0770	+0613	+0856	+0755	+0855	+0677	+0784	+0784	+0784	+0784	+0784	+0784

Use of Prediction XV

It is necessary to prepare a tape containing the operators and a tape containing the traces x_1, y_1, z_1, u_1. These are prepared as described in the following two pages. Assume these are given tape numbers X and Y respectively. Then the operating instruction would be:

Erase storage, put S11 switch off

Read in 2539 m 2

Read in 2539 P —— (Control Tape)

Read in X

Place Y in Photoelectric Reader

Start at 145

The control tapes control the output and serve the following functions:

2539 - P0	Print errors and sums of squares and scope display sums of squares	DIRECT PRINTOUT
2539 - P1	Print sums of squares and scope display sums of squares	
2539 - P2	Scope display sums of squares	
2539 - P3	Print errors and sums of squares	
2539 - P4	Print sums of squares	
2539 - P5	Print nothing, display nothing	
2539 - P10	Print errors and sums of squares and scope display sums of squares	DELAYED PRINTOUT (MAGNETIC TAPE)
2539 - P11	Print sums of squares and scope display sums of squares	
2539 - P12	Print errors and sums of squares	
2539 - P13	Print sums of squares	

If one of the operators on X were badly prepared, it might happen that machine overflow would occur causing the machine to stop while computing for that operator. If this does happen, starting the machine over at 166 will have the effect of ignoring the bad operator and proceeding to the remaining ones.

Preparation of Data Parameter

Each set of data x_i, y_i, z_i, and u_i is prepared as a separate parameter and then the four parameters are combined into one long one. The form of each is identical.

Octal Address		Octal Address		Octal Address		Octal Address	
1054	$\bar{x}$	1054	$\bar{y}$	1054	$\bar{z}$	1054	$\bar{u}$
1055	x_1	1055	y_1	1055	z_1	1054	u_1
1056	x_2	1056	y_2	1056	z_2	1056	u_2
↓	↓	↓	↓	↓	↓	↓	↓
2037	x_{499}	2037	y_{499}	2037	z_{499}	2037	u_{499}
Start at 1033		Start at 1033		Start at 1033		Start at 1033	

Notes:

It is not necessary that the series contain 499 members. However, there must be four traces. If less than four are to be used, short dummy traces must be inserted. For consistency with the operator tape, the order of combination of the separate parameters must be x_i, y_i, z_i, u_i.

The data must appear as integers in the range -99 through +99.

Preparation of Operator Parameters

Up to 8 operators may be prepared on a single tape in the following fashion.

Octal Address	Contents	Explanation		
1053	+N	N = no. of operators on tape		
1054	+.XXXX	Ident. no. for first operator		
1055	-0,-1,-2 or -3	-0 if x_1, -1 if y_1		
1056	+k	-2 if z_1, -3 if u_1		
1057	+M			
1060	±.XXXX	$a \times 10^{-3}$		1rst OPERATOR
1061	±.XXXX	$a_7 \times 10^{-1}$	Constants for x_1	
1062	±.XXXX	$a_6 \times 10^{-1}$		
.	.	.		
.	.	.		
1070	.	$a_0 \times 10^{-1}$		
1071	.	$b_7 \times 10^{-1}$	Constants for y_1	
.	.	.		
1100	.	$b_0 \times 10^{-1}$		
1101	.	$c_7 \times 10^{-1}$	Constants for z_1	
.	.	.		
1110	.	$c_0 \times 10^{-1}$		
1111	.	$d_7 \times 10^{-1}$	Constants for u_1	
.	.	.		
1120	.	$d_0 \times 10^{-1}$		
1121	.	Ident. no. for second operator		2nd OP.
.	.	etc.		
.	.			
1165	.			
:	.	etc.		

Notes:

It is not necessary to put anything into irrelevant registers. For example, if the first operator had an M of 3 registers 1061-1064, 1071-1074, 1101-1104, and 1121-1124 would be considered irrelevant. Again, if this operator did not use the u_1 series in its prediction mechanism, registers 1111-1120 would be irrelevant.

SHEET 1

Address	Contents	Remarks
00		
01		
02		
03		
04		
05		
06		
07		
10		
11		
12		
13		
14		
15		
16		
17		
20		
21		
22		
23		
24		
25		
26		
27		
30		
31		
32		
33		
34		
35		
36		
37		
40	ta 64	
41	ta 72	
42	ta 115	
43	sp 116	
44	ta 112	
45	cm 115	
46	mh 73	
47	ad 74	
50	td 53	
51	sr 17	
52	ta 115	
53	ca	
54	rc 54	
55	sp 135	
56	ao 112	
57	cp 45	
60	ca 107	
61	rc 61	
62	sp 135	
63	ao 114	
64	cp	
65	ca 110	
66	rc 66	
67	sp 135	
70	ca 113	
71	ta 114	
72	sp	
73	+10	
74	0.00075	
75	1.74000	
76	0.52000	
77	0.36000	

NOTES

Address	Contents	Remarks
100	0.16000	
101	0.26000	
102	0.46000	
103	0.66000	
104	0.56000	
105	0.06000	
106	1.54000	
107	0.20000	
110	1.22000	
111	+3	
112	si 0	PRINT
113	si 0	
114	si 0	
115	si 0	
116	cp 126	
117	ca 124	
120	rc 120	
121	sp 135	
122	ca 111	
123	sp 44	
124	0.32000	
125	0.72000	
126	ca 125	
127	sp 120	
130	ta 134	
131	ca 142	
132	ad 144	
133	cp 132	
134	sp	
135	ta 141	
136	ca 143	
137	ad 144	
140	cp 137	
141	sp	
142	-140	
143	-128	
144	+1	
145	sp 550	PROGRAM CONTROL
146	sp 1002	
147	sp 1027	
150	sp 1011	
151	sp 234	
152	sp 262	
153	sp 617	
154	sp 356	
155	sp 441	
156	sp 476	
157	sp 332	
160	sp 306	
161	sp 617	
162	sp 372	
163	sp 455	
164	sp 476	
165	sp 344	
166	sp 607	
167	si 0	
170	si 0	
171	ca 224	E (LAST HALF)
172	td 204	
173	td 207	
174	ca 225	
175	td 217	
176	ca 226	
177	ta 227	

Address	Contents	Remarks
200	ca 230	FORM $\Sigma 10^{x}$ IN l_{20} THRU l_{43}
201	ta 231	
202	ca 0	
203	ta 232	
204	ca	
205	mh 233	
206	sr 17	
207	mh	
210	ad 232	
211	ta 232	
212	ao 204	
213	ao 207	
214	ao 231	
215	cp 204	
216	ca 232	
217	ta	
220	ao 217	
221	ao 227	
222	cp 200	
223	sp	
224	ca 1460	
225	ca 1100	
226	-23	
227	si 0	
230	-9	
231	si 0	
232	si 0	
233	+10	
234	ca 256	B1 OPERATORS IN
235	si 703	
236	ca 257	
237	ti 1007	
240	ca 256	
241	ad 257	
242	ta 256	
243	sp 152	
244	ca 260	
245	ta 261	
246	ca 224	
247	td 251	
250	ca 1013	
251	ta	
252	ao 251	
253	ao 261	
254	cp 250	
255	sp 657	
256	+4097	
257	+37	
260	-239	
261	si 0	
262	ca 301	B2 F/2 TRACES IN
263	si 703	
264	ca 300	
265	ti 1054	
266	ca 301	
267	ad 305	
270	ta 301	
271	ao 303	
272	cp 277	
273	ca 302	
274	ta 301	
275	ca 304	
276	ta 303	
277	sp 153	

Address	Contents	Remarks
300	+260	
301	+2049	
302	+2049	
303	-3	
304	-3	
305	+500	
306	ca 325	B3 S/2 TRACES IN
307	si 703	
310	ca 324	
311	ti 1054	
312	ca 325	
313	ad 331	
314	ta 325	
315	ao 327	
316	cp 323	
317	ca 326	
320	ta 325	
321	ca 330	
322	ta 327	
323	sp 161	
324	+260	
325	+2289	
326	+2289	
327	-3	
330	-3	
331	+500	
332	ca 343	A3 F/2 ERRORS OUT
333	si 707	
334	ca 342	
335	ta 1460	
336	ca 343	
337	ad 342	
340	ta 343	
341	sp 160	
342	+240	
343	+6145	
344	ca 355	A4 S/2 ERRORS OUT
345	si 707	
346	ca 354	
347	ta 1460	
350	ca 355	
351	ad 354	
352	ta 355	
353	sp 166	
354	+240	
355	+8193	
356	ta 431	F1 F2 PRINT ERRORS
357	si 225	
360	sp 130	
361	ca 110	
362	rc 362	
363	sp 135	
364	ca 1007	
365	sp 40	
366	ca 110	
367	rc 367	
370	sp 135	
371	sp 375	
372	ta 431	
373	si 225	
374	sp 130	
375	ca 436	
376	ta 113	
377	ta 114	

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE PREDICTION XV INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 2539 m2

DL-481

400	ca 224	AND SUMS OF ERRORS SQUARED ↓
401	td 410	
402	ca 225	
403	td 423	
404	ca 434	
405	ts 435	
406	ca 432	
407	ts 433	
410	ca	
411	sp 40	
412	ao 410	
413	ao 433	
414	cp 410	
415	ca 107	
416	rc 416	
417	sp 135	
420	ca 107	
421	rc 421	
422	sp 135	
423	ca	
424	sp 40	
425	ao 423	
426	ao 435	
427	cp 406	
430	si 630	
431	sp	
432	1-9	
433	si 0	
434	-23	
435	si 0	
436	-10	
437	1.77747	
440	1.77747	
441	ta 475	G1 G2 PRINT SUMS OF ERRORS SQUARED
442	si 225	
443	sp 130	
444	ca 110	
445	rc 445	
446	sp 135	
447	ca 1007	
450	sp 40	
451	ca 110	
452	rc 452	
453	sp 135	
454	sp 460	
455	ta 475	
456	si 225	
457	sp 130	
460	ca 434	
461	ts 113	
462	ts 114	
463	ca 225	
464	td 467	
465	ca 434	
466	ts 435	
467	ca	
470	sp 40	
471	ao 467	
472	ao 435	
473	cp 467	
474	si 630	
475	sp	
476	ta 541	H
477	ca 225	

NOTES

500	td 532	H SCOPE DISPLAY SUMS OF ERRORS SQUARED
501	ca 434	
502	ts 435	
503	sp 774	
504	td 440	
505	ca 542	
506	ts 543	
507	ca 437	
510	mh 547	
511	slr 12	
512	si 600	
513	ca 543	
514	rc 514	
515	ca 543	
516	ad 544	
517	ts 543	
520	rc 520	
521	su 545	
522	cp 515	
523	ao 440	
524	ao 440	
525	mh 547	
526	slr 12	
527	si 600	
530	ca 542	
531	rc 531	
532	ca	
533	slr 1	
534	rc 534	
535	ao 532	
536	ao 435	
537	cp 523	
540	si 500	
541	sp	
542	1.57777	
543	1.57777	
544	0.06315	
545	0.60000	
546	0.20000	
547	+.6667	
550	ca 1053	C RESET PROGRAM CONTROL
551	ts 170	
552	ao 170	
553	ao 1047	
554	cp 563	
555	ca 606	
556	td 357	
557	td 373	
560	td 442	
561	td 456	
562	td 611	
563	ao 1050	
564	cp 574	
565	ao 563	
566	ao 574	
567	ao 575	
570	ao 571	
571	ao 605	
572	cp 563	
573	sp 146	
574	ca 601	
575	td 154	
576	ad 604	
577	td 162	

600	sp 565	
601	sp 155	
602	sp 156	
603	sp 157	
604	+6	
605	-2	
606	si 136	
607	rc 170	J PROG END
610	cp 151	
611	si 0	
612	ca 616	
613	rc 613	
614	si 630	
615	si 0	
616	1.42000	
617	ta 223	
620	ts 746	E
621	ca 746	
622	su 1001	
623	ts 746	
624	sp 625	
625	ca 1010	
626	ts 1003	
627	ca 760	
630	td 670	FORM ERRORS
631	ca 761	
632	td 702	
633	su 1011	
634	su 1012	
635	ts 763	
636	ca 762	
637	td 704	
640	td 705	
641	ca 754	
642	su 1012	
643	ts 764	
644	ca 755	i=20
645	su 1012	
646	ts 765	TO
647	ca 756	
650	su 1012	
651	ts 766	i=259
652	ca 757	
653	su 1012	AND
654	ts 767	
655	sp 244	
656	si 0	
657	ca 1012	i=260
660	ts 772	
661	ca 1012	TO
662	ts 773	
663	ao 773	
664	ca 330	i=499
665	ts 1002	
666	ca 763	
667	td 721	
670	ca	
671	td 722	
672	ca 762	
673	td 723	
674	td 724	
675	ao 670	
676	ac 1003	
677	cp 715	

SHEET 2

700	ca 771	
701	ts 1004	
702	ca	
703	mr 770	
704	ad	
705	ts	
706	ao 702	ERRORS LEFT IN REGISTERS 1460 THRU 2037
707	ao 704	
710	ao 705	
711	ao 1004	
712	cp 702	
713	ca 233	
714	ts 1013	
715	ca 771	
716	ts 1004	
717	ca 1012	
720	ts 1005	
721	ca	
722	mr	
723	ad	
724	ts	
725	ao 721	
726	ao 722	
727	ao 1005	
730	cp 721	
731	ca 721	
732	ad 777	
733	ts 721	
734	ca 722	
735	su 773	
736	ts 722	
737	ao 723	
740	ao 724	
741	ao 1004	
742	cp 717	
743	ca 752	
744	td 624	
745	ao 1002	
746	cp	
747	ca 753	
750	td 624	
751	sp 771	
752	sp 666	
753	sp 625	
754	ca 1023	
755	ca 1033	
756	ca 1043	
757	ca 1053	
760	ca 764	
761	ca 1100	
762	ca 1460	
763	si 0	
764	si 0	
765	si 0	
766	si 0	
767	si 0	
770	+.1000	
771	-239	
772	si 0	
773	si 0	
774	si 500	
775	ca 437	
776	sp 504	
777	si 0	

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE PREDICTION XV INDEX
AUTHOR SIMPSON DATE
TAPE NUMBER 2539 m 2

DL-481

SHEET 3

1000	si 0			
1001	+2			
1002	ca 1007	A1		
1003	si 707			
1004	ca 1010			
1005	to 1053			
1006	sp 147			OPERATORS
1007	+4096			T
1010	+277			-0 -1 -2 -3
1011	ca 1025	A2		+K
1012	si 707			+M
1013	ca 1074		T	$a \times 10^{-3}$
1014	to 1054		R	$a_7 \times 10^{1}$
1015	ca 1025		A	a_6
1016	ad 1024		C	a_5
1017	ts 1025		E	a_4
1020	ao 1026		S	a_3
1021	cp 1023			a_2
1022	sp 151		O	a_1
1023	sp 147		U	$a_0 \times 10^{-1}$
1024	+500		T	b_7
1025	+2049			b_6
1026	-3			b_5
1027	sp 25	D		b_4
1030	+.0100		T	b_3
1031	ts 2037		R	b_2
1032	ca 1054		A C	b_1
1033	ca 1032		E S	$b_0 \times 10^{-1}$
1034	td 1036		I	c_7
1035	td 1041		N	c_6
1036	ca		S	c_5
1037	mh 1030		C	c_4
1040	sr 17		A	c_3
1041	ts		L E	c_2
1042	ar 1036		F	c_1
1043	ao 1041		A	$c_0 \times 10^{-1}$
1044	su 1031		C T	d_7
1045	cp 1036		O R	d_6
1046	sp 150			d_5
1047	si 0			d_4
1050	si 0			d_3
1051	si 0			d_2
1052	si 0			d_1
1053	si 0			$d_0 \times 10^{-1}$
1054				
1055				
1056				
1057				
1060				
1061				
1062				
1063				
1064			O	
1065			P	
1066			E	
1067			R	
1070			A	
1071			T	
1072			O	
1073			R	
1074				
1075			O	
1076			N	
1077			E	

1100		
1101		
1102		
1103		
1104		
1105		
1106		
1107		
1110		
1111		
1112		
1113		
1114		
1115		
1116		
1117		
1120		
1121		
1122		
1123		
1124		
1125		
1126		
1127		T
1130		W
1131		O
1132		
1133		
1134		
1135		
1136		
1137		
1140		
1141		
1142		
1143		
1144		
1145		
1146		
1147		
1150		
1151		
1152		
1153		
1154		
1155		
1156		
1157		
1160		
1161		
1162		
1163		
1164		
1165		
1166		
1167		
1170		E
1171		T
1172		C
1173		
1174		
1175		
1176		
1177		

NOTES

00		00	
01		01	
02		02	
03		03	
04		04	
05		05	
06		06	
07		07	
10		10	
11		11	
12		12	
13		13	
14		14	
15		15	
16		16	
17		17	
20		20	
21		21	
22		22	
23		23	
24		24	
25		25	
26		26	
27		27	
30		30	
31		31	
32		32	
33		33	
34		34	
35		35	
36		36	
37		37	
40		40	
41		41	
42		42	
43		43	
44		44	
45		45	
46		46	
47		47	
50		50	
51		51	
52		52	
53		53	
54		54	
55		55	
56		56	
57		57	
60		60	
61		61	
62		62	
63		63	
64		64	
65		65	
66		66	
67		67	
70		70	
71		71	
72		72	
73		73	
74		74	
75		75	
76		76	
77		77	

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE PREDICTION XV INDEX
AUTHOR SIMPSON DATE
TAPE NUMBER 2539 m 2

DL-461

APPENDIX C

APPENDIX C

ITERATION I (2615 m 2) - Description and Use Of

This program was written with the purpose of obtaining least squares fits for linear operators as described in Part IV. It computes essentially as described, but has provisions for changing its increment after cycling for a certain prescribed number of times.

The program was designed to be run in conjunction with the Prediction XV program described in Appendix B, and to illustrate the conveniences which programs can include. The data to which the linear operator is to be fitted is prepared in the same fashion as in Prediction XV. The information about the operators to be found (Iteration I solves up to eight operators one after the other) is prepared as a single tape. The operator coefficients once formed are printed out, and also are left in the machine in a form to be used directly with the prediction program.

The output of Iteration I was designed to eliminate identification problems. In addition to printing out the coefficients identified, it prints out the operator number, the operator parameters including which set of data is predicted, and the variance and minimum sums of square errors. These last two numbers allow a rapid computation of the percent reduction $R = 1 - I_{min}/I_{var}$, which is a measure of the goodness of the least squares fit. A sample of the output appears below.

```
Variance sum = 004953
Minimum sum  = 001840
Operator No. -1010
N = 066
n = 050
k = 002
M = 003
T4 predicted
a/1000 = +0292

a3/10  = -0000
a2/10  = +0565
a1/10  = -0000
a0/10  = +0005

d3/10  = -0341
d2/10  = -0117
d1/10  = -0410
d0/10  = +0708
```

The program may be used to print out all the values of I as they are computed. A plot of these values for a particular operator appears in Part IV.

One other feature in this program is a "roll back" procedure. This permits us to avoid having to start from scratch if the machine fails in the middle of the long computations. Every fifteen seconds during the computation, all of electrostatic storage is transferred to the magnetic drums which are very reliable. Then if electrostatic storage is destroyed, we can call back the program from the magnetic drums and start over where we left off not more than fifteen seconds ago.

Use of Iteration I

If the operators are prepared as described on the next page with a tape number X then the instructions for the operation of this program would be:

Erase storage, S11 switch down

Read in 2615 m 2

Read in 2615 P (control tape)

Read in Y (data tape)

Place X in photoelectric reader

Start over at 145

The control tapes control the output and serve the following functions:

2615 P 0 Print out operator and identification (direct printer)

2615 P 1 Print out operator, identification, and all values of I (delayed printer).

Preparation of Operator Parameters

The information for each operator is prepared as a short separate tape and the tapes are then combined in any order. The form of each operator is identical.

Octal Address	Contents	Explanation
1001	+N	First member op. interval
1002	+n	Length of op. interval
1003	+0, or -1	-1 if x_1 not used
1004	+0, or -1	-1 if y_1 " "
1005	+0, or -1	-1 if z_1 " "
1006	+0, or -1	-1 if u_1 " "
1007	+.XXXX	First operator no.
1010	$\pm$0,-1,-2,or -3	-0 if x_1, -1 if y_1
1111	+k	-2 if z_1, -3 if u_1
1112	+M	
1113	$\pm$.XXXX	= mean of pred. series x 10^{-3}

Start at 147

The first of the separate tapes must have one additional register, register 1000, which contains + no. of operators on the combined tape.

Register 2123 contains +.0010 which is the first increment to be used for the a term. Register 2223 contains +.0100 which is the first increment to be used with the remaining constants. Register 3421 is the counter for the cycles at these increments. The second set of increments is 1/10 the first set, and appears in registers 2124 and 2224. The counter for this set is register 3433. These registers may be changed to adapt to the particular problem.

The "roll back" procedure in case of electrostatic storage failure is

Erase storage

Read in 2615 P 13

Start over at 145

SHEET 1

00	
01	
02	
03	
04	
05	
06	
07	
10	
11	
12	
13	
14	
15	
16	
17	
20	
21	
22	
23	
24	
25	
26	
27	
30	
31	
32	
33	
34	
35	
36	
37	
40	ta 64
41	ta 72
42	ta 115
43	sp 116
44	ta 112
45	cm 115
46	mh 73
47	ad 74
50	td 53
51	sh 17
52	ta 115
53	ca
54	rc 54
55	sp 135
56	ao 112
57	cp 45
60	ca 107
61	rc 61
62	sp 135
63	ao 114
64	cp
65	ca 110
66	rc 66
67	sp 135
70	ca 113
71	ta 114
72	sp
73	+10
74	0.00075
75	1.74000
76	0.52000
77	0 36000

100	0.16000
101	0.26000
102	0.46000
103	0.66000
104	0.56000
105	0.06000
106	1.58000
107	0.20000
110	1.22000
111	+3
112	si 0
113	si 0
114	si 0
115	si 0
116	cp 126
117	ca 124
120	rc 120
121	sp 135
122	ca 111
123	sp 44
124	0.32000
125	0.72000
126	ca 125
127	sp 120
130	td 134
131	ca 142
132	ad 144
133	cp 132
134	sp
135	td 141
136	ca 143
137	ad 144
140	cp 137
141	sp
142	-140
143	-128
144	+1
145	sp 146
146	sp 25
147	sp 3232
150	sp 261
151	sp 617
152	sp 2614
153	sp 2100
154	sp 2160
155	sp 2204
156	sp 2227
157	sp 2252
160	sp 3400
161	sp 3425
162	sp 163
163	sp 2670
164	sp 2737
165	sp 2376
166	sp 2550
167	sp 2560
170	sp 2570
171	sp 2600
172	sp 225
173	sp 600
174	sp 217
175	si 0
176	si 0
177	si 0

NOTES

200	ca 215
201	si 703
202	ad 202
203	ta 214
204	ao 214
205	sp 207
206	si 0
207	ca 215
210	si 707
211	ca 214
212	rc 212
213	sp 173
214	si 0
215	+4096
216	si 0
217	si 0
220	ca 224
221	rc 221
222	si 630
223	si 0
224	1 42000
225	si 630
226	ca 237
227	si 707
230	ca 240
231	bo 1007
232	sa 237
233	ad 240
234	ta 237
235	sp 200
236	si 0
237	+4097
240	+37
241	si 0
242	si 0
243	ca 256
244	ta 257
245	ca 255
246	td 250
247	ca 1013
250	ta
251	ao 250
252	ao 257
253	cp 247
254	sp 657
255	ca 1460
256	si 0
257	si 0
260	si 0
261	ta 277
262	ca 301
263	si 703
264	ca 300
265	bi 1054
266	ca 301
267	ad 305
270	ta 301
271	ao 303
272	cp 277
273	ta 302
274	ta 301
275	ca 304
276	ta 303
277	sp

300	si 0
301	si 0
302	si 0
303	-3
304	-3
305	+500
306	si 0
307	si 0
310	ca 255
311	td 320
312	td 323
313	ca 344
314	ta 343
315	ca 0
316	ta 341
317	ta 342
320	ca
321	mh 345
322	sh 17
323	mh
324	mh 345
325	sh 17
326	sa 342
327	ta 342
330	ca 341
331	ta 341
332	ao 320
333	ao 323
334	ao 343
335	cp 320
336	sp 451
337	sp
340	si 0
341	si 0
342	si 0
343	si 0
344	si 0
345	+10
346	si 0
347	ta 424
350	ca 74
351	td 413
352	td 416
353	ca 404
354	td 365
355	ca 402
356	ta 403
357	ca 401
360	td 370
361	ca 341
362	ta 407
363	ao 407
364	ca 407
365	su
366	cp 372
367	ta 407
370	ao
371	sp 364
372	ao 365
373	ao 370
374	ao 370
375	ao 370
376	ao 403
377	cp 364

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION I INDEX ______
AUTHOR SIMPSON DATE ______
TAPE NUMBER 2615-m 0

DL-481

SHEET 2

400	sp 411
401	ca 413
402	-1
403	si 0
404	ca 405
405	+10
406	+1
407	si 0
410	si 0
411	si 225
412	sp 130
413	ca
414	rc 414
415	sp 135
416	ca
417	rc 417
420	sp 135
421	ca 342
422	sp 40
423	si 630
424	sp
425	si 0
426	si 0
427	ta 467
430	sp 434
431	ta 474
432	ca 342
433	ta 475
434	ca 476
435	ta 473
436	ao 473
437	su 474
440	cp 457
441	su 406
442	cp 444
443	sp 465
444	ca 477
445	su 475
446	cp 457
447	sp 465
450	si 0
451	ca 341
452	ta 474
453	ca 342
454	ta 475
455	sp 337
456	si 0
457	ca 462
460	ts 500
461	sp 467
462	+0
463	si 0
464	si 0
465	ca 470
466	ts 500
467	sp
470	-1
471	+2
472	si 0
473	si 0
474	si 0
475	si 0
476	si 0
477	si 0

500	si 0
501	si 0
502	si 0
503	si 0
504	si 0
505	ca 513
506	si 707
507	ca 514
510	bo 40
511	sp 516
512	si 0
513	+6145
514	+1857
515	si 0
516	ao 1047
517	cp 535
520	sp 522
521	si 136
522	ca 521
523	td 217
524	td 411
525	td 2361
526	td 2376
527	td 2614
530	td 2670
531	sp 535
532	si 0
533	si 0
534	si 0
535	ao 1050
536	cp 540
537	sp 562
540	ca 567
541	ts 2114
542	ca 570
543	ts 2137
544	ca 571
545	ts 2147
546	ca 572
547	ts 2170
550	ca 573
551	ts 2214
552	ca 574
553	ts 2237
554	ca 575
555	ts 2262
556	ca 576
557	ts 2314
560	ca 577
561	ts 2337
562	ca 604
563	si 707
564	ca 0
565	rc 565
566	si 0
567	sp 2115
570	sp 2140
571	sp 2150
572	sp 2171
573	sp 2215
574	sp 2240
575	sp 2263
576	sp 2315
577	sp 2340

NOTES

600	ao 177
601	cp 146
602	sp 174
603	si 0
604	+4096
605	si 0
606	si 0
607	+10
610	-3
611	+2
612	si 0
613	si 0
614	si 0
615	si 0
616	si 0
617	ta 337
620	ta 746
621	ca 146
622	su 611
623	ts 746
624	sp 625
625	ca 1010
626	ts 613
627	ca 760
630	td 670
631	ca 761
632	td 702
633	su 1011
634	su 1012
635	ts 763
636	ca 762
637	td 764
640	td 705
641	ca 754
642	su 1012
643	ts 764
644	ca 755
645	su 1012
646	ts 765
647	ca 756
650	su 1012
651	ts 766
652	ca 757
653	su 1012
654	ts 767
655	sp 243
656	si 0
657	ca 1012
660	ts 772
661	ca 1012
662	ts 773
663	ao 773
664	ca 610
665	ts 612
666	ca 763
667	td 121
670	ca
671	td 722
672	ca 162
673	td 723
674	td 124
675	ao 610
676	ao 613
677	cp 715

700	ca 771
701	ts 614
702	ca
703	mr 770
704	ad
705	ts
706	ao 702
707	ao 704
710	ao 705
711	ao 614
712	cp 702
713	ca 607
714	ts 613
715	ca 771
716	ts 614
717	cs 1012
720	ts 615
721	ca
722	mr
723	ad
724	ts
725	ao 721
726	ao 722
727	ao 615
730	cp 721
731	ca 721
732	ad 772
733	ts 721
734	ca 722
735	su 773
736	ts 722
737	ao 723
740	ao 724
741	ao 614
742	cp 717
743	ca 752
744	td 624
745	ao 612
746	cp
747	ca 753
750	td 624
751	sp 310
752	sp 666
753	sp 625
754	ca 1023
755	ca 1033
756	ca 1043
757	ca 1053
760	ca 764
761	ca 1100
762	ca 1460
763	si 0
764	si 0
765	si 0
766	si 0
767	si 0
770	+.1000
771	si 0
772	si 0
773	si 0
774	si 0
775	si 0
776	si 0
777	si 0

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION I INDEX ____
AUTHOR SIMPSON DATE ____
TAPE NUMBER 2615-m0

DL-481

SHEET 3

1000	si 0	3500	sl 225	3600	td 3531	00	
1001	si 0	3501	ca 2756	3601	sp 3350	01	
1002	si 0	3502	rc 3502	3602	si 0	02	
1003	si 0	3503	sp 135	3603	td 2355	03	
1004	si 0	3504	do 3512	3604	td 3553	04	
1005	si 0	3505	cp 3501	3605	sp 3355	05	
1006	si 0	3506	ca 3513	3606	si 0	06	
1007	si 0	3507	ta 3512	3607	si 0	07	
1010	si 0	3510	sp 150	3610	ca 2355	10	
1011	ca 1025	3511	si 0	3611	ta 3553	11	
1012	si 707	3512	qp 3772	3612	sp 2300	12	
1013	ca 1024	3513	qp 3772	3613	si 0	13	
1014	to 1054	3514	si 0	3614	si 0	14	
1015	ca 1025	3515	si 0	3615	ta 3621	15	
1016	ad 1024	3516	si 0	3616	ao 3552	16	
1017	ts 1025	3517	si 0	3617	ao 3554	17	
1020	ao 1026	3520	ca 513	3620	sp 2277	20	
1021	ca 1023	3521	si 707	3621	sp 0	21	
1022	si 0	3522	ca 514	3622	si 0	22	
1023	sp 1027	3523	to 40	3623	si 0	23	
1024	+500	3524	sp 2100	24		24	
1025	+2049	3525	si 0	25		25	
1026	-3	3526	si 0	26		26	
1027	sp 25	3527	si 0	27		27	
1030	+.0100	3530	ca 1013	30		30	
1031	ta 2031	3531	su 2123	31		31	
1032	ca 1054	3532	ta 1013	32		32	
1033	ca 1032	3533	ca 2126	33		33	
1034	td 1036	3534	ta 3531	34		34	
1035	td 1041	3535	ca 476	35		35	
1036	ca	3536	ta 474	36		36	
1037	mr 1030	3537	ca 477	37		37	
1040	slr 17	3540	ta 475	40		2040	qp 3774
1041	ta	3541	sp 154	41		2041	ca 3452
1042	ao 1036	3542	si 0	42		2042	rc 2042
1043	ao 1041	3543	si 0	43		2043	sp 135
1044	su 1031	3544	si 0	44		2044	do 2041
1045	cp 1036	3545	si 0	45		2045	ao 2040
1046	sp 1011	3546	si 0	46		2046	cp 2041
1047	si 0	3547	si 0	47		2047	ca 3456
1050	si 0	3550	ca 2355	50		2050	ta 2040
1051	si 0	3551	ta 2333	51		2051	ca 2076
1052	si 0	3552	ca 0	52		2052	td 2041
1053	si 0	3553	su 2322	53		2053	ca 1010
1054		3554	ta 0	54		2054	ta 3451
1055		3555	ca 2326	55		2055	ao 3451
1056		3556	ta 3553	56		2056	ad 74
1057		3557	ca 476	57		2057	td 2060
1060		3560	ta 474	60		2060	ca 0
1061		3561	ca 477	61		2061	rc 2061
1062		3562	ta 475	62		2062	sp 135
1063		3563	sp 2347	63		2063	ca 3460
1064		3564	si 0	64		2064	rc 2064
1065		3565	ta 3571	65		2065	sp 135
1066		3566	td 2334	66		2066	do 2063
1067		3567	td 3552	67		2067	ao 3475
1070		3570	td 3554	70		2070	cp 2063
1071		3571	sp 0	71		2071	ca 3476
1072		3572	do 3531	72		2072	ta 3475
1073		3573	ao 3553	73		2073	ca 2077
1074		3574	sp 161	74		2074	td 2063
1075		3575	si 0	75		2075	sp 165
1076		3576	si 0	76		2076	ca 3452
1077		3577	td 2155	77		2077	ca 3460

NOTES.

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION 1 INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 2615-m0

DL-481

SHEET 4

2100	ca 474	2200	sp 3615
2101	ex 476	2201	sp 2170
2102	ca 475	2202	si 0
2103	ex 477	2203	si 0
2104	sp 2105	2204	ca 1012
2105	ca 1013	2205	ts 2202
2106	ad 2123	2206	ca 2226
2107	ts 1013	2207	td 2305
2110	ca 2127	2210	td 2307
2111	td 2104	2211	td 2332
2112	sp 261	2212	sp 3565
2113	sp 617	2213	sp 2277
2114	sp 347	2214	sp 2360
2115	sp 427	2215	ao 2202
2116	ao 500	2216	cp 2220
2117	cp 2100	2217	sp 156
2120	ca 2126	2220	ao 2307
2121	ts 2133	2221	ao 2332
2122	sp 2151	2222	ao 2334
2123	+.0010	2223	ao 2305
2124	+.0001	2224	sp 3615
2125	+4	2225	sp 2214
2126	su 2123	2226	si 0
2127	sp 2132	2227	ca 1012
2130	si 0	2230	ts 2202
2131	si 0	2231	ca 2251
2132	ca 1013	2232	td 2305
2133	ad 2123	2233	td 2307
2134	ts 1013	2234	td 2332
2135	sp 261	2235	sp 3565
2136	sp 617	2236	sp 2277
2137	sp 347	2237	sp 2360
2140	sp 427	2240	ao 2202
2141	ao 500	2241	cp 2243
2142	cp 2100	2242	sp 157
2143	ca 2154	2243	ao 2307
2144	td 2104	2244	ao 2332
2145	ca 2155	2245	ao 2334
2146	ts 2133	2246	ao 2305
2147	sp 2360	2247	sp 3615
2150	sp 3530	2250	sp 2237
2151	ca 2155	2251	si 0
2152	ts 3531	2252	ca 1012
2153	sp 2100	2253	ts 2202
2154	sp 2105	2254	ca 2274
2155	ad 2123	2255	td 2305
2156	si 0	2256	td 2307
2157	si 0	2257	td 2332
2160	ca 1012	2260	sp 3565
2161	ts 2202	2261	sp 2277
2162	ca 2203	2262	sp 2360
2163	td 2305	2263	ao 2202
2164	td 2307	2264	cp 2266
2165	td 2332	2265	sp 160
2166	sp 3565	2266	ao 2307
2167	sp 2277	2267	ao 2332
2170	sp 2360	2270	ao 2334
2171	ao 2202	2271	ao 2305
2172	cp 2174	2272	sp 3615
2173	sp 155	2273	sp 2262
2174	ao 2307	2274	si 0
2175	ao 2332	2275	si 0
2176	ao 2334	2276	si 0
2177	ao 2305	2277	ta 2347

NOTES

2300	ca 474	2400	ca 2421
2301	ex 476	2401	rc 2401
2302	ca 475	2402	sp 135
2303	ex 477	2403	ao 2400
2304	sp 2305	2404	ao 2437
2305	ca 0	2405	cp 2400
2306	ad 2323	2406	ca 2440
2307	ts 0	2407	ts 2437
2310	ca 2327	2410	ca 2441
2311	td 2304	2411	td 2400
2312	sp 261	2412	ca 2442
2313	sp 617	2413	ts 113
2314	sp 347	2414	ts 114
2315	sp 427	2415	ca 1013
2316	ao 500	2416	sp 40
2317	cp 2300	2417	ca 2423
2320	ca 2326	2420	rc 2420
2321	ts 2333	2421	sp 135
2322	sp 3610	2422	sp 166
2323	+.0100	2423	1.22000
2324	+.0010	2424	0.14000
2325	+.0001	2425	1.62000
2326	su 2323	2426	0.52000
2327	sp 2332	2427	1.72000
2330	si 0	2430	0.52000
2331	si 0	2431	1.74000
2332	ca 0	2432	1.74000
2333	ad 2323	2433	1.74000
2334	ts 0	2434	0.20000
2335	sp 261	2435	0.22000
2336	sp 617	2436	0.20000
2337	sp 347	2437	-11
2340	sp 427	2440	-11
2341	ao 500	2441	ca 2423
2342	cp 2300	2442	-100
2343	ca 2354	2443	si 0
2344	td 2304	2444	ca 2465
2345	sp 3550	2445	si 225
2346	ts 2333	2446	ca 2465
2347	sp 0	2447	rc 2447
2350	mh 770	2450	sp 135
2351	sl 13	2451	ao 2446
2352	mh 770	2452	ao 2437
2353	sp 1040	2453	cp 2446
2354	sp 2305	2454	ca 2440
2355	ad 2323	2455	ts 2437
2356	si 0	2456	ca 2444
2357	si 0	2457	td 2446
2360	ta 2373	2460	ca 0
2361	si 2373	2461	sp 40
2362	ca 11[illegible]	2462	ca 2423
2363	rc 2363	2463	rc 2463
2364	sp 135	2464	sp 2410
2365	ca 110	2465	1.22000
2366	rc 2366	2466	si 0
2367	sp 135	2467	si 0
2370	si 630	2470	1.62000
2371	ca 114	2471	0.32000
2372	ts 113	2472	1.72000
2373	sp 0	2473	0.52000
2374	si 0	2474	1.14000
2375	si 0	2475	[illegible]
2376	si [illegible]	2476	0.20000
2377	sp 13[illegible]	2477	0.2[illegible]000

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER [illegible]15 m c

DL-481

2500	0.20000	2600	ca 2541
2501	ta 2523	2601	td 2505
2502	ca 2465	2602	ca 2543
2503	rc 2503	2603	td 2507
2504	sp 135	2604	sp 2501
2505	ca 0	2605	sp 172
2506	ta 2466	2606	si 0
2507	ca 0	2607	si 0
2510	td 2460	2610	sp 135
2511	ca 74	2611	sp 2521
2512	ad 1012	2612	si 0
2513	td 2516	2613	si 0
2514	ca 1012	2614	si 225
2515	ta 2535	2615	ca 2644
2516	ca 0	2616	rc 2616
2517	ta 2467	2617	sp 135
2520	sp 2445	2620	ac 2615
2521	ao 2535	2621	ac 2641
2522	cp 2524	2622	cp 2615
2523	sp 0	2623	ca 2642
2524	ab 2460	2624	ta 2641
2525	ca 2516	2625	ca 2640
2526	su 2546	2626	td 2615
2527	ta 2516	2627	sp 247
2530	sp 2516	2630	ca 2644
2531	0.14000	2631	rc 2721
2532	1.44000	2632	sp 135
2533	0.70000	2633	ca 2666
2534	0.44000	2634	ta 113
2535	si 0	2635	ta 114
2536	ca 2531	2636	si 630
2537	ca 2532	2637	sp 153
2540	ca 2533	2640	ca 2644
2541	ca 2534	2641	-17
2542	ca 764	2642	17
2543	ca 765	2643	si 0
2544	ca 766	2644	1.22000
2545	ca 767	2645	1.62000
2546	-1	2646	1.70000
2547	si 0	2647	1.72000
2550	ca 2536	2650	0.14000
2551	td 2505	2651	0.50000
2552	ca 2542	2652	0.30000
2553	td 2507	2653	0.14000
2554	sp 2501	2654	0.60000
2555	sp 167	2655	0.70000
2556	si 0	2656	0.04000
2557	si 0	2657	0.20000
2560	ca 2537	2660	0.24000
2561	td 2505	2661	0.34000
2562	ca 2543	2662	1.60000
2563	td 2507	2663	0.20000
2564	sp 2501	2664	0.22000
2565	sp 170	2665	0.20000
2566	si 0	2666	9
2567	si 0	2667	si 0
2570	ca 2540	2670	si 225
2571	td 2505	2671	ca 2712
2572	ca 2544	2672	rc 2672
2573	td 2507	2673	sp 135
2574	sp 2501	2674	ac 2641
2575	sp 171	2675	ac 2641
2576	si 0	2676	cp 2671
2577	si 0	2677	ca 2642

NOTES

SHEET 5

2700	[illegible] 64	3000	[illegible] 2756
2701	ca 2[illegible]	01	si 0
2702	td [illegible]	02	si 0
03	sp 34[illegible]	03	ca 2644
04	ca 2644	04	rc 3004
2705	rc 2705	05	sp 135
2706	sp 135	06	ca 2644
2707	ca 2442	07	rc 3007
2710	ta 114	10	sp 135
2711	sp 164	11	ca 3030
2712	1.22000	12	[illegible]
2713	1.62000	13	sp 135
2714	1.60000	14	ac 3011
2715	1.72000	15	[illegible] 3036
2716	0.30000	16	cp 3011
2717	0.60000	17	ca 3037
2720	0.[illegible]0000	20	ta 3036
2721	1.[illegible]0000	21	ca 3040
2722	0.34000	22	td 3011
2723	1.6[illegible]000	23	ca 1001
24	0.20000	24	sp 3050
25	0.24000	25	sp 3042
26	0.24000	26	si 0
27	1.70000	27	si 0
30	0.20000	30	1.62000
31	0.20000	31	0.60000
32	0.22000	32	1.72000
33	0.[illegible]0000	33	0.[illegible]0000
34	ca [illegible]712	34	0.22000
2735	ta 114	35	0.[illegible]0000
36	si 225	36	-5
37	ca 2756	37	5
40	rc 2740	40	ca 3030
41	sp 135	41	si 0
42	ac 2737	42	ca 3066
43	ac 2641	43	rc [illegible] 43
44	cp 2724	44	sp 135
45	ca 2642	45	ac 3042
46	td 2641	46	ac 3046
47	ca 3000	47	sp 3042
50	td 2727	50	ca 3066
51	ca 1007	51	ta 3065
52	sp 40	52	ca 3067
53	sp 3003	53	td 3042
54	si [illegible]	54	ca 1002
55	si 0	55	sp 3150
56	1.22000	56	sp 3071
57	1.[illegible]000	57	si 0
60	1.40000	60	1.22000
61	1.[illegible]000	61	[illegible]0000
62	1.30000	62	0.20000
63	0.04000	63	0.12000
64	0.50000	64	0.20000
65	0.14000	65	-4
66	1.00000	66	-4
67	1.40000	67	ca 3060
70	0.[illegible]0000	70	si 0
71	[illegible]	71	ca 3107
72	1.62000	72	rc 3072
73	0.60000	73	sp 135
74	1.72000	74	ac 3071
75	1.[illegible]000	75	ac 3065
76	0.72000	76	cp 3071
77	0.10000	77	ca 3066

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION I INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 2615 - m.c

DL-481

SHEET 6

Loc	Instr	Loc	Instr	Loc	Instr	Loc	Instr
3100	ta 3065	3200	sp 3165	3300	td 166	3400	ao 3421
3101	ca 3114	3201	sp 3215	3301	ao 1004	3401	cp 153
3102	td 3071	3202	si 0	3302	cp 3304	3402	ca 3422
3103	ca 1011	3203	0.00075	3303	sp 3310	3403	ts 3421
3104	sp 3150	3204	ca 3215	3304	ca 3332	3404	ca 3420
3105	sp 3116	3205	-2	3305	td 155	3405	td 160
3106	si 0	3206	si 0	3306	ad 3330	3406	ao 2106
3107	1.22000	3207	ca 3210	3307	td 167	3407	ao 2126
3110	0.74000	3210	+100	3310	ao 1005	3410	ao 2133
3111	0.20000	3211	+10	3311	cp 3313	3411	ao 2155
3112	0.22000	3212	+1	3312	sp 3317	3412	ao 2306
3113	0.20000	3213	si 0	3313	ca 3333	3413	ao 2326
3114	ca 3107	3214	si 0	3314	td 156	3414	ao 2333
3115	si 0	3215	ca 0	3315	ad 3330	3415	ao 2355
3116	ca 3134	3216	rc 3216	3316	td 170	3416	sp 3572
3117	rc 3117	3217	sp 135	3317	ao 1006	3417	si 0
3120	sp 135	3220	ca 0	3320	cp 3322	3420	sp 161
3121	ao 3116	3221	rc 3221	3321	sp 3326	3421	-2
3122	ao 3143	3222	sp 135	3322	ca 3334	3422	-2
3123	cp 3116	3223	ca 0	3323	td 157	3423	si 0
3124	ca 3144	3224	rc 3224	3324	ad 3330	3424	si 0
3125	ts 3143	3225	sp 135	3325	td 171	3425	ao 3433
3126	ca 3145	3226	sp 0	3326	sp 3343	3426	cp 153
3127	td 3116	3227	si 0	3327	si 0	3427	ca 3434
3130	ca 1012	3230	si 0	3330	+10	3430	ts 3433
3131	sp 3152	3231	si 0	3331	sp 155	3431	sp 162
3132	sp 2041	3232	sp 3235	3332	sp 156	3432	si 0
3133	si 0	3233	ca 3241	3333	sp 157	3433	-1
3134	1.22000	3234	si 701	3334	sp 160	3434	-1
3135	1.62000	3235	ca 1000	3335	+2047	3435	+.1000
3136	1.60000	3236	rc 3236	3336	+20	3436	-36
3137	1.72000	3237	sp 3242	3337	si 0	3437	-36
3140	0.20000	3240	si [illegible]	3340	ca 2123	3440	ca 3437
3141	0.22000	3241	+614[illegible]	3341	ca 2223	3441	ts 3436
3142	0.20000	3242	ca 1000	3342	si 0	3442	ca 0
3143	-6	3243	ts 176	3343	ca 3340	3443	ts 0
3144	-6	3244	cs 1000	3344	td 2106	3444	ao 3443
3145	ca 3134	3245	ts 177	3345	td 2126	3445	ao 3436
3146	si 0	3246	ac 177	3346	td 2133	3446	cp 3442
3147	si 0	3247	ca 3252	3347	sp 3557	3447	sp 3500
3150	ta 3226	3250	td 232	3350	ca 3341	3450	ts 3474
3151	ta 3213	3251	sp 3253	3351	td 2306	3451	si 0
3152	ar 3213	3252	sp 3253	3352	td 2326	3452	cm 2000
3153	ca 3203	3253	ca [illegible]	3353	td 2353	3453	sr 2000
3154	td 3215	3254	ad 1001	3354	sp 3603	3454	ca 0
3155	td 3220	3255	su 3226	3355	ca 754	3455	cl 2000
3156	td 3223	3256	ts 302	3356	su 1012	3456	qp 3774
3157	ca 3207	3257	ts 301	3357	td 2203	3457	si 0
3160	td 3[illegible]	3260	ca 1002	3360	ca 755	3460	bo 0
3161	ca 3205	3261	ad 3336	3361	su 1012	3461	ao 0
3162	ta 3206	3262	ts 300	3362	td 2226	3462	ta 0
3163	ca 3204	3263	ca 1002	3363	ca 756	3463	ra 0
3164	td 3171	3264	ts 356	3364	su 1012	3464	td 0
3165	ca 3213	3265	ts 344	3365	td 2251	3465	qh 0
3166	su 0	3266	ts 111	3366	ca 757	3466	cp 0
3167	sp 3173	3267	ac 56	3367	su 1012	3467	ck 0
3170	ta 3213	3270	ac 244	3370	td 2274	3470	ra 0
3171	ac 0	3271	ac 771	3371	ca 3317	3471	td 0
3172	sp 3165	3272	ac 1002	3372	td 160	3472	cm 2000
3173	su 3166	3273	cp 3215	3373	ca 3276	3473	cm 2000
3174	ac 3171	3274	sp 3301	3374	td 3442	3474	si 0
3175	rc 3171	3275	ca 3331	3375	sp 3440	3475	qp 3766
3176	ac 2171	3276	td 154	3376	ca 1014	3476	qp 3766
3177	rc 3206	3277	ad 3330	3377	sp 3400	3477	si 0

NOTES

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE ITERATION I INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 2615 - m. c

DL-481

APPENDIX D

APPENDIX D

AUTO CROSS-CORRELATION I 2559 mo, ml) - Description and Use Of

This program was written for the WWI Digital Computer to compute the unnormalized sample correlations

$$\sum_{i+N}^{N+n-1} x_{i-j} y_i \qquad j = 0, 1, 2 \ldots . m$$

The conventions for preparation of the data x_i and y_i are identical with those described for Prediction XV in Appendix B, with the exception that the data tapes need not be combined after preparation. A short tape is prepared containing the information N, n, and m as follows

Register	1047	+ N	(First data point in block)
(Octal)	1050	+ n	(No. data points in block)
	1051	+ m	(No. lags)

2559 mo handles individual data tapes and is used as follows

Erase storage, put Sil down

Read in 2559 mo

Read in Z (tape for N, n, m)

Read in X (x_i data tape)

Read in Y (y_i data tape)

Start at 770

If X and Y are not identical we get half of the cross-correlation curve (for j 0). To get the other half, we repeat the instructions interchanging the order of read-in for X and Y. If X and Y are the same tape, we get the entire auto-correlation curve, since since auto-correlations are symmetric about the zeroth lag.

The correlations are printed out by the direct printer as seven-place numbers, ten per line, the 0th lag being the first no. on the first line, the 1rst lag being the second no. on the first line, etc.

2559 ml performs the same functions as 2559 mo, but is adapted for handling the combined tapes used with Prediction XV. It assumes there are 3 real data sets plus a dummy set and forms the nine correlations representing the permutations of the 3 real sets. The correlations are over 380 values of the data, and are taken to 100 lags. The output is the delayed printer, and requires one minute for each 100 lag block. At this rate the program can perform 8 or 10 million multiplications in 4 hours of machine time. A sample of the output is shown on the next page.

RECORD 12.4 T1, T3, T5

CORRELATIONS = n' $Q'_{T--T--}(p)$

N' = 120 m = 100

n' = 370 p = 0, -1, ... -m

T1T1

0179270 0175549 0167943 0159926 0154392 0152791 0154375 0157500 0160762 0163416
0165321 0166495 0166739 0165643 0163303 0160803 0159622 0160656 0163452 0166588
0168334 0167537 0164736 0161344 0158992 0158675 0159932 0161713 0162974 0163369
0163530 0164083 0164976 0165711 0165652 0164657 0163094 0161890 0161770 0162461
0163614 0163930 0162907 0161306 0160078 0160104 0161483 0163441 0164974 0165438
0165304 0165441 0166164 0166961 0166724 0164631 0161027 0157575 0156120 0157333
0160723 0164614 0167174 0167842 0166975 0165494 0164158 0163148 0162496 0161983
0161894 0162752 0164481 0166457 0167456 0166564 0163950 0160785 0158852 0158994
0161202 0164385 0166921 0167969 0167684 0166672 0165477 0164313 0163157 0161878
0160869 0160636 0161380 0162796 0164121 0164902 0164818 0164335 0164288 0165130
0166570

T1T3

0212315 0205051 0197189 0192194 0191950 0195871 0201416 0206027 0208399 0208603
0207664 0206130 0204026 0201293 0198671 0197364 0198592 0202105 0206145 0208921
0209045 0206355 0202538 0199672 0198995 0200283 0202215 0203689 0204039 0203696
0203808 0204576 0205619 0206097 0205229 0203193 0200919 0199747 0200230 0201889
0203855 0204841 0204594 0203746 0203076 0203147 0203796 0204298 0204135 0203438
0203065 0203750 0205317 0206341 0205592 0202834 0199241 0196915 0197372 0200464
0204846 0208429 0209584 0208257 0205568 0203089 0201738 0201436 0201732 0202350
0203410 0205086 0206897 0207946 0207196 0204530 0201016 0198505 0198493 0201004
0205076 0208558 0209854 0208729 0206316 0204030 0202591 0202043 0202011 0201972
0202154 0202876 0204135 0205599 0206456 0206337 0205253 0204088 0203818 0204721
0206194

T1T5

0209619 0202418 0194992 0190346 0189852 0192921 0197627 0202017 0204816 0205892
0205566 0203982 0201230 0197769 0194739 0193592 0195427 0199749 0204563 0207811
0208031 0205093 0200592 0196696 0195019 0195593 0197376 0199289 0200417 0200991
0201573 0202362 0202995 0202941 0202052 0200418 0198843 0198474 0199452 0200979
0202027 0201708 0200246 0198710 0198077 0198724 0200240 0201682 0202178 0201819
0201532 0202012 0202953 0203148 0201688 0198684 0195532 0194158 0195609 0199468
0203951 0206958 0207149 0204766 0201330 0198496 0197102 0197084 0197926 0199340
0201398 0204089 0206618 0207553 0205910 0201907 0197105 0193901 0193894 0197030
0201804 0205869 0207635 0206750 0204291 0201551 0199462 0198337 0198049 0198426
0199550 0201315 0203328 0204797 0204898 0203555 0201258 0199341 0198886 0200234
0202612

T3T1

0212315 0214230 0211085 0204997 0198972 0195694 0195736 0198067 0201243 0204156
0206571 0208142 0208394 0206896 0203669 0199515 0196011 0195022 0197272 0201639
0206213 0208676 0208284 0205975 0203138 0201366 0201268 0202300 0203388 0203614
0203158 0202595 0202514 0202965 0203421 0203311 0202552 0201699 0201672 0202735
0204740 0206443 0206737 0205565 0203606 0202126 0201797 0202280 0202770 0202386
0201275 0200503 0201133 0203442 0206422 0208297 0207773 0205107 0202023 0200134
0200408 0202296 0204504 0205787 0205749 0204874 0203941 0203295 0202763 0201748
0200511 0200046 0201303 0204326 0207785 0209951 0209440 0206467 0202834 0200394
0200202 0202080 0204417 0205863 0205834 0204829 0203754 0203166 0203132 0203255
0203320 0203525 0204280 0205588 0206847 0207156 0205938 0203269 0200563 0199272
0200207

T3T3

0269909 0265706 0256802 0247574 0241891 0241348 0244756 0249524 0253665 0256073
0257163 0257240 0256225 0253851 0250495 0247405 0246542 0248938 0253606 0258193
0260357 0258809 0254777 0250492 0247962 0248053 0249871 0251853 0252758 0252534
0252496 0253346 0255092 0256991 0257643 0256624 0254305 0251877 0250637 0250896
0252183 0253218 0253426 0253018 0252728 0253066 0253749 0254090 0253453 0251917
0250751 0251120 0253476 0256785 0259049 0258778 0255780 0251673 0248717 0248300
0250421 0253577 0255923 0256450 0255513 0254166 0253326 0253018 0252810 0252240
0251881 0252629 0254885 0257828 0259654 0259015 0255674 0251246 0248299 0248297
0251198 0255194 0257873 0258140 0256306 0254031 0252778 0252825 0253670 0254324
0254635 0254841 0255504 0256390 0256789 0255943 0253700 0250774 0248926 0249346
0252154

T3T5

0261579 0258137 0251129 0244258 0240413 0240513 0243623 0247978 0252005 0254770
0256106 0255702 0253348 0249351 0244530 0240709 0239840 0242870 0248441 0254093
0257400 0256909 0253511 0249110 0245885 0245181 0246616 0248910 0250438 0250744
0250528 0250280 0250387 0250502 0250142 0249157 0247968 0247496 0248343 0250252
0252304 0253160 0252283 0250366 0248668 0248153 0248641 0249393 0249538 0248787
0247998 0248275 0250144 0252699 0254218 0253440 0250483 0247015 0245271 0246195
0249339 0252849 0254797 0254377 0252184 0249797 0248130 0247402 0247085 0246813
0247203 0249075 0252563 0256210 0257999 0256650 0252239 0247033 0243746 0244026
0247474 0252038 0255261 0255942 0254281 0251630 0249399 0248193 0247942 0248265
0248955 0250277 0252326 0254405 0255303 0254251 0251263 0247568 0245103 0245277
0248127

T5T6

0209619 0212038 0209335 0203211 0196693 0192461 0191581 0193231 0196093 0199052
0201607 0203523 0204436 0204000 0201893 0198661 0195687 0194688 0196313 0199712
0203166 0204746 0203810 0201116 0198049 0196123 0196040 0197551 0198966 0199887
0200298 0200580 0201113 0201808 0202130 0201613 0200460 0199281 0198848 0199346
0200502 0201601 0201728 0200788 0199293 0198161 0197832 0198239 0198926 0199360
0199702 0200359 0201613 0203350 0204737 0204639 0202457 0198811 0195347 0193572
0194388 0197391 0201058 0203921 0205094 0204666 0203335 0201642 0199916 0198151
0196876 0196723 0198201 0201054 0204070 0205587 0204639 0201459 0197718 0195364
0195317 0197459 0200387 0202653 0203503 0203103 0202116 0201188 0200492 0199752
0199000 0198644 0199082 0200357 0201726 0202234 0201376 0199169 0196832 0195679
0196311

T5T3

0261579 0258503 0251130 0243226 0238177 0237759 0241323 0246520 0251191 0253999
0255065 0254873 0253629 0251155 0247622 0243998 0241989 0242889 0246389 0250736
0253821 0254032 0251731 0248511 0246101 0245658 0246846 0248696 0249412 0249165
0248910 0249316 0250478 0251667 0251882 0250681 0248649 0246845 0246399 0247486
0249367 0250990 0251434 0250840 0249792 0249027 0248655 0248462 0248069 0247487
0247406 0248440 0250518 0252857 0253862 0252575 0249266 0245444 0243258 0243786
0246985 0251304 0254515 0255452 0254075 0251484 0249071 0247474 0246661 0246353
0246782 0248200 0250730 0253666 0255346 0254600 0251279 0246782 0243491 0243226
0246072 0250554 0253998 0254889 0253248 0250477 0248352 0247568 0247855 0248337
0248540 0248758 0249401 0250568 0251588 0251501 0249919 0247283 0245280 0245305
0247641

T5T5

0260952 0257123 0248990 0240358 0234742 0233724 0236575 0241565 0246583 0250383
0252531 0252789 0251067 0247427 0242673 0238511 0236926 0239032 0243935 0249400
0253017 0253106 0249976 0245377 0241455 0239781 0240557 0243023 0245084 0246490
0247433 0248105 0248531 0248305 0247163 0245313 0243479 0242682 0243616 0245862
0248378 0249779 0249186 0247196 0244753 0243055 0242447 0242861 0243792 0244802
0246006 0247554 0249418 0250780 0250442 0247878 0243718 0239863 0238360 0240083
0244495 0249677 0253109 0253525 0251174 0247524 0244207 0242064 0241157 0241370
0242851 0245681 0249543 0253041 0254352 0252348 0247416 0241608 0237715 0237649
0241224 0246595 0250978 0252717 0251670 0248923 0246000 0243937 0242980 0242918
0243585 0245071 0247265 0249468 0250540 0249533 0246532 0242624 0239773 0239506
0241958

Use of 2559 ml

The instructions for use are

Erase storage, S11 switch down

Read in 2559 ml

Read in W (combined data)

Start at 677

If the combined tape has 4 real data sets, and we want the 16 permutations of correlation, then an additional tape is used and the instructions are

Erase storage, S11 switch down

Read in 2559 ml

Read in 2559 p4

Read in W

Start at 677

2559 ml is equipped with the same "roll back" procedure that Iteration I is (Appendix C). In case of machine failure

Erase storage

Read in 2559 p13

Start at 677

Traveling Correlations

With the aid of tape 2559 p10 we can use 2559 mo to obtain correlations from highly overlapping blocks of the data. The correlations are over blocks 50 in length and the number of lags is taken to be 20. The first reading in each block has an index (N) equal to k x 10 where k = 3, 4, ... 44. The procedure for using 2559 mo in this way is

Erase storage, put Sil down

Read in 2559 mo

Read in X

Read in Y

Read in 2559 p10   leave in P.E.T.R.

Start at 770   (21 lags are printed for N = 30)

Read in

Start at 770   (21 lags are printed for N = 40)

Read in

Start at 770   (21 lags are printed for N = 50)

etc.

Start at 770   (21 lags are printed for N = 440)

[SHEET 1]

	00	ca 1051		
	1001	ta 1023		
	1002	sp 734		
	1003	si 0		
	1004	si 0		
	1005	si 0		
	1006	si 0		
	1007	si 0		C
	1010	si 0		O
	1011	si 0		N
	1012	+500		S
	1013	+1		T
	1014	si 0		A
	1015	si 0		N
	1016	si 0		T
	1017	si 0		S
	1020	si 0		
	1021	ca 2030		
	1022	ca 1054		
	1023	si 0		
	1024	si 0		
	1025	si 0		
	1026	si 0		
	1027	si 0		
	30			
	31			
	32			
	33			
	34			
	35			
	36			
	37			
	40	ta 63		
	41	ta 70		
	42	si 225		
	43	ta 113		
	44	cs 107		
	45	ta 110		
	46	cm 113		
	47	mh 71		
	50	ad 72		
	51	td 54		
	52	slr 17		P
	53	ta 113		R
	54	ca		I
	55	rc 55		N
	56	ao 110		T
	57	cp 46		
	60	ca 105		R
	61	rc 61		O
	62	ao 112		U
	63	cp		T
	64	ca 106		I
	65	rc 65		N
	66	ca 111		E
	67	ta 112		
	70	sp		
	71	+10		
	72	0.00073		
	73	1.74000		
	74	0.52000		
	75	0.36000		
	76	0.16000		
	77	0.26000		

	100	0.46000		
	101	0.66000		
	102	1.56000		
	103	0.06000		
	104	1.54000		
	105	0.20000		
	106	122000		
	107	+3		
	110	si 0		
	111	-9		
	112	-9		
	113	si 0		
	14			
	15			
	16			
	17			
	20			
	21			
	22			
	23			
	24			
	25			
	26			
	27			
	1030	+.0100		
	1031	ts 2031		
	1032	ca 1054		S
	1033	ca 1032		C
	1034	td 1036		A
	1035	td 1041		L
	1036	ca		E
	1037	mh 1030		
	1040	slr 17		F
	1041	ts		A
	1042	ao 1036		C
	1043	ao 1041		T
	1044	su 1031		O
	1045	cp 1036		R
	1046	sp 560		Y
	1047		N	
	1050		# CROSS PRODUCTS	
	1051		# LAGS	
	52			
	53			
	54			
	55			
	56			
	57			
	560	ca 1032		
	561	td 566		M
	562	ca 603		O
	563	ts 604		V
	564	sp 566		E
	565	ao 567		
	566	ca		T
	567	ts 3014		R
	570	ao 566		A
	571	ao 604		C
	572	cp 565		E
	573	ca 567		S
	574	td 605		
	575	ca 605		
	576	ad 606		
	577	td 567		

NOTES

	600	si 0		
	601	si 0		
	602	si 0		C
	603	-499		O
	604	si 0		N
	605	si 0		S
	606	0.06027		T
	607	si 0		A
	610	si 0		N
	611	si 0		T
	612	si 0		S
	613	si 0		
	614	si 0		
	615	si 0		
	616	si 0		
	617	ca 650		
	620	td 663		
	621	td 665		
	622	td 667		
	623	ca 654		
	624	td 634		
	625	ca 652		
	626	ts 652		
	627	ca 651		
	630	td 637		O
	631	ao 722		V
	632	ts 660		E
	633	ca 660		R
	634	su		F
	635	cp 641		L
	636	ts 660		O
	637	ao		W
	640	sp 633		
	641	ao 634		P
	642	ao 637		R
	643	ao 637		I
	644	ao 653		N
	645	cp 633		T
	646	sp 662		
	647	si 0		
	650	0.00013		
	651	ca 663		
	652	-2		
	653	si 0		
	654	ca 655		
	655	+100		
	656	+10		
	657	+1		
	660	si 0		
	661	si 0		
	662	si 225		
	663	ca		
	664	rc 664		
	665	ca		
	666	rc 666		
	667	ca		
	670	rc 670		
	671	ca 723		
	672	sp 40		
	673	ao 1023		
	674	cp 734		
	675	si 0	STOP	
	676	si 0		
	677	si 0		

	700	ca 1032		
	701	td 711		
	702	ca 0		
	703	ts 722		S
	704	ts 723		U
	705	ca 1050		M
	706	ts 724		
	707	ao 724		P
	710	ca 723		R
	711	sa		O
	712	ts 723		D
	713	ca 722		U
	714	ts 722		C
	715	ac 711		T
	716	ao 724		S
	717	cp 710		
	720	sp 617		
	721	si 0		
	722	si 0		
	723	si 0		
	724	si 0		
	725	si 0		
	726	ca 1015		
	727	su 1013		
	730	ts 1015		
	731	sp 700		
	732	si 0		
	733	si 0		
	734	ca 1050		
	735	ts 1024		
	736	ao 1024		
	737	ca 1022		
	740	td 753		
	741	ca 1016		F
	742	td 752		O
	743	ca 1015		R
	744	td 751		M
	745	sp 751		
	746	si 0		P
	747	si 0		R
	750	si 0		O
	751	ca		D
	52	mr		U
	753	ts		C
	754	ca 751		T
	755	su 1013		S
	756	ts 751		
	757	ca 752		
	760	su 1013		
	761	ts 752		
	762	ao 753		
	763	ao 1024		
	764	cp 751		
	765	sp 726		
	766	si 0		
	767	si 0		
START	770	ca 1021		
	771	ad 1047		
	772	ts 1014		S
	773	ad 1054		E
	774	su 1013		T
	775	ts 1015		
	776	ad 1012		
	777	ts 1016		

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE AUTO CROSS-CORRELATION INDEX ________

AUTHOR SIMPSON DATE ________

TAPE NUMBER 2559-m0

DL-481

300	ca 313	500	ao 507
301	si 707	501	cp 320
302	ca 312	502	si 176
303	ta 1054	503	ca 510
304	ca 313	504	rc 504
305	ad 312	505	si 134
306	ts 313	506	si 0
307	ao 314	507	-8
310	cp 677	510	1.42000
311	sp 320	11	
312	+500	12	
313	+2049	13	
314	-3	14	
15		15	
16		16	
17		17	
320	ca 334	20	
321	si 703	21	
322	ca 312	22	
323	ti 2030	23	
324	ao 335	24	
325	cp 340	25	
326	ts 334	26	
327	ad 312	27	
330	ts 334	30	
331	ca 336	31	
332	ts 335	32	
333	sp 340	33	
334	+2049	34	
335	-2	35	
336	-2	36	
37		37	
340	ca 356	540	ca 111
341	si 713	541	ts 112
342	ca 312	542	si 176
343	ti 3014	543	ca 106
344	ca 356	544	rc 544
345	ad 312	545	ts 554
346	ts 356	546	cp 543
347	ao 360	547	ca 553
350	cp 770	550	ts 554
351	ca 336	551	si 134
352	ts 360	552	sp 500
353	ca 357	553	-4
354	ts 356	554	-4
355	sp 770	55	
356	+2049	56	
357	+2049	57	
360	-2	60	
61		61	
62		662	si 176
63		63	
64		64	
65		65	
66		66	
67		67	
370	ca 376	70	
371	si 707	71	
372	ca 375	72	
373	ta 40	673	si 134
374	sp 320	674	su 1023
375	d.01054	675	cp 134
376	+4097	676	sp 540
77		677	sp 32

NOTES.

DL-481

SHEET 2

00		00	
01		01	
02		02	
03		03	
04		04	
05		05	
06		06	
07		07	
10		10	
11		11	
12		12	
13		13	
14		14	
15		15	
16		16	
17		17	
20		20	
21		21	
22		22	
23		23	
24		24	
25		25	
26		26	
27		27	
30		30	
31		31	
32		32	
33		33	
34		34	
35		35	
36		36	
37		37	
40		40	
41		41	
42	sp 43	42	
43		43	
44		44	
45		45	
1046	sp 300	46	
1047	+120	47	
1050	+370	50	
1051	+100	51	
52		52	
53		53	
54		54	
55		55	
56		56	
57		57	
60		60	
61		61	
62		62	
63		63	
64		64	
65		65	
66		66	
67		67	
70		70	
71		71	
72		72	
73		73	
74		74	
75		75	
76		76	
77		77	

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM

TITLE AUTO CROSS-CORRELATION INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 6559 m 1

APPENDIX E

APPENDIX E

BIOGRAPHICAL NOTE Stephen Milton Simpson, Jr.

Attended Yale University September 1946 - June 1950, receiving B.S. in Physics. Entered the Massachusetts Institute of Technology in the Department of Geology in September 1950. Member of Phi Beta Kappa and Sigma Xi.

Presently under appointment as Instructor in the Department of Geology and Geophysics at the Massachusetts Institute of Technology.