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
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 require-
ments for the degree of Doctor of Philosophy.

Several specific problems in seismic and gravi-
tational 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 investi-
gated 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 compar-
sion of these contours with each other (for various order poly-
nomials) 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
generically. 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 autocorrelations, cross-correlations and "traveling" auto- or cross-correlations.

Thesis Supervisor: Patrick M. Hurley
Title: Professor of Geology
# TABLE OF CONTENTS

Acknowledgments .......................................................... 1
Introduction ................................................................... 11

## PART I LEAST SQUARES RESIDUAL GRAVITY

Introduction .................................................................. I-1
Theory .......................................................................... I-2
Simplified Solutions ....................................................... I-5
  Case n = 1 .................................................................. I-6
  Case n = 2 .................................................................. I-7
  Case n = 3 .................................................................. I-8
  Case n = 4 .................................................................. I-11
Discussion ..................................................................... I-15
Applications ................................................................... I-16
Example ....................................................................... I-17
Setting up the Normal Equations ................................. I-20

## PART II SEISMIC RECORD ANALYSIS BY LINEAR OPERATORS

Introduction .................................................................. II-1
Statistical Methods ......................................................... II-2
Single Frequency Cosine Operators ............................. II-4
Geometry of Cosine Operators ...................................... II-6
Conclusions .................................................................. II-8
Example ....................................................................... II-9
Cyclical Nature of Cosine Operators ............................. II-10
Cosine Operator Predicting an Autoregressive Series .... II-11
Cosine Operators on Series with Gaussian Spectrum Distribution ................................................ II-14
Computational Example ................................................ II-15
A Method of Finding Linear Operators for Least Squares Fitting Procedures ................................ II-18
Accuracy ....................................................................... II-22

## PART III SOME INTERPRETIVE PROCEDURES

Introduction .................................................................. III-1
Velocity Separation ......................................................... III-2
Tests of the Method ........................................................ III-6
Conclusions .................................................................. III-9
Phase Test .................................................................... III-10
Ensemble Average .......................................................... III-13
Travelling Correlations .................................................. III-15

Conclusions and Suggestions for Further Work

References

APPENDIX A Polynomial I Description and Use Of
APPENDIX B Prediction XV Description and Use Of
APPENDIX C Iteration I Description and Use Of
APPENDIX D Auto Cross-Correlation I Description and Use Of
APPENDIX E Biographical Note
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 artificial 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 \(y\).

\[
G(xy) = \sum_{i=0}^{n} \sum_{j=0}^{n-1} c_{ij} x^i y^j
\]

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

\[
\sum_{xy} R^2(xy) = \sum_{xy} \left[ \sum_{i=0}^{n} \sum_{j=0}^{n-1} c_{ij} x^i y^j \sum_{k=0}^{n} \sum_{l=0}^{n-k} c_{kl} x^k y^l \right]
\]

\[
-2 \sum_{xy} g(xy) \sum_{i=0}^{n} \sum_{j=0}^{n-1} c_{ij} x^i y^j + \sum_{xy} g^2(xy)
\]
or
\[ \sum_{xy} R^2(xy) = \sum_{xy} \left[ \sum_{i=0}^{n} \sum_{j=0}^{n-i} \sum_{k=0}^{n} \sum_{l=0}^{n-k} c_{ij} c_{kl} x^{k+i} y^{l+j} \right] \]
\[ -2 \sum_{xy} \left[ \sum_{i=0}^{n} \sum_{j=0}^{n-i} c_{ij} x^{i} y^{j} \right] + \sum_{xy} g^2(xy) \]

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

\[ \sum_{k=0}^{n} \sum_{l=0}^{n-k} c_{kl} \sum_{xy} x^{k+i} y^{l+j} = \sum_{xy} g(xy) x^{i} y^{j} \quad 1.2 \]

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

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 \( c_{k'l'} \)'s. Once these two are chosen, a unique inverse matrix exists, which, if found, may be used to compute the \( c_{k'l'} \)'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.
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 abscissae 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 $c_{k,l}$'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 $c_{k,l}$'s in terms of the summations $\Sigma g(xy)x^i y^j$. 

To demonstrate how this is done we shall solve the equations for $n = 1, 2, 3, \text{ and } 4$, 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:

- $\Sigma x^0 y^0 = M = 121$
- $\Sigma x^2 = \Sigma y^2 = 1210$
- $\Sigma x^4 = \Sigma y^4 = 21538$
- $\Sigma x^6 = \Sigma y^6 = 451330$
- $\Sigma x^8 = \Sigma y^8 = 10185538$
- $\Sigma x^2 y^2 = \Sigma x^4 y^2 = 12100$
- $\Sigma x^4 y^4 = \Sigma x^6 y^2 = 215380$
- $\Sigma x^6 y^6 = \Sigma x^8 y^2 = 4513300$
- $\Sigma x^8 y^4 = \Sigma y^8 = 3833764$

† Ref. 2
Case $n = 1$

The normal equations are

$$
\begin{align*}
c_{00}M + c_{10}Ex + c_{01}Ey &= \Sigma g(xy) \\
c_{00}Ex + c_{10}Ex^2 + c_{01}Exy &= \Sigma g(xy)x \\
c_{00}Ey + c_{10}Exy + c_{01}Ey^2 &= \Sigma g(xy)y
\end{align*}
$$

Reducing immediately to

$$
\begin{align*}
c_{00} = \frac{\Sigma g(xy)}{M} & \quad c_{10} = \frac{\Sigma g(xy)x}{\Sigma x^2} & \quad c_{01} = \frac{\Sigma g(xy)y}{\Sigma y^2}
\end{align*}
$$

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

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

Case $n = 2$

The normal equations are

$$
\begin{align*}
\begin{array}{ccccccc}
c_{00} & c_{01} & c_{02} & c_{11} & c_{10} & c_{20} \\
N & \Sigma y & \Sigma y^2 & \Sigma xy & \Sigma x & \Sigma x^2 & = \Sigma g(xy) \\
\Sigma y & \Sigma y^2 & \Sigma y^3 & \Sigma xy^2 & \Sigma xy & \Sigma x^2y & = \Sigma g(xy)y \\
\Sigma y^2 & \Sigma y^3 & \Sigma y^4 & \Sigma xy^3 & \Sigma xy^2 & \Sigma x^2y^2 & = \Sigma g(xy)y^2 \\
\Sigma xy & \Sigma xy^2 & \Sigma xy^3 & \Sigma x^2y & \Sigma x^2y^2 & \Sigma xy^3 & = \Sigma g(xy)xy \\
\Sigma x & \Sigma xy & \Sigma xy^2 & \Sigma x^2y & \Sigma x^3 & \Sigma x^3 & = \Sigma g(xy)x \\
\Sigma x^2 & \Sigma x^2y & \Sigma x^2y^2 & \Sigma x^3y & \Sigma x^3 & \Sigma x^4 & = \Sigma g(xy)x^2
\end{array}
\end{align*}
$$

which reduce to

$$
\begin{align*}
c_{01} &= \frac{1}{1210} \Sigma g(xy)y & c_{11} &= \frac{1}{12100} \Sigma g(xy)xy & c_{10} &= \frac{1}{1210} \Sigma g(xy)x
\end{align*}
$$

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

I-6
$1210_{00} + 1210_{02} + 1210_{20} = \Sigma g(xy)$

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

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

The solutions are

\[
\begin{align*}
0_{00} & = \frac{1}{9438} [278\Sigma g(xy) - 10(\Sigma g(xy)x^2 + \Sigma g(xy)y^2)] \\
0_{02} & = \frac{1}{9438} [\Sigma g(xy)y^2 - 10\Sigma g(xy)] \\
0_{20} & = \frac{1}{9438} [\Sigma g(xy)x^2 - 10\Sigma g(xy)]
\end{align*}
\]

Thus

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

\[
+ y[\frac{1}{1210}\Sigma g(xy)y]
\]

\[
+ y^2[\frac{1}{9438} \Sigma g(xy)y^2 - 10\Sigma g(xy)]
\]

\[
+ xy[\frac{1}{12100}\Sigma g(xy)xy]
\]

\[
+ x[\frac{1}{1210}\Sigma g(xy)x]
\]

\[
+ x^2[\frac{1}{9438} \Sigma g(xy)x^2 - 10\Sigma g(xy)]
\]

Or to six places

\[
G(xy) = [0.0294554\Sigma g(xy) - 1.05955 \times 10^{-3}(\Sigma g(xy)x^2 + \Sigma g(xy)y^2)]
\]

\[
+ y[8.26448 \times 10^{-4}\Sigma g(xy)y]
\]

\[
+ y^2[1.05955 \times 10^{-4}(\Sigma g(xy)y^2 - 10\Sigma g(xy))]
\]

\[
+ xy[8.26448 \times 10^{-5}\Sigma g(xy)xy]
\]

\[
+ x[8.26448 \times 10^{-4}\Sigma g(xy)x]
\]

\[
+ x^2[1.05955 \times 10^{-4}(\Sigma g(xy)x^2 - 10\Sigma g(xy))]
\]
Case \( n = 3 \)

The normal equations are

\[
\begin{array}{cccccccc}
0_00 & 0_01 & 0_02 & 0_03 & 0_11 & 0_12 & 0_10 & 0_20 & 0_30 \\
1) & M & y & y^2 & y^3 & xy & xy^2 & x^2 & x^3 = g \\
2) & y & y^2 & y^3 & y^4 & xy^2 & xy^3 & x^2y & x^3y = gy \\
3) & y^2 & y^3 & y^4 & y^5 & xy^3 & xy^4 & x^2y^2 & x^3y^2 = gy^2 \\
4) & y^3 & y^4 & y^5 & y^6 & xy^4 & xy^5 & x^2y^3 & x^3y^3 = gy^3 \\
5) & xy & xy^2 & xy^3 & xy^4 & x^2y^2 & x^2y^3 & x^3y^2 & x^3y^3 = gxy \\
6) & xy^2 & xy^3 & xy^4 & xy^5 & x^2y^3 & x^2y^4 & x^3y^3 & x^3y^4 = gxy^2 \\
7) & x^2y & x^2y^2 & x^2y^3 & x^2y^4 & x^2y^5 & x^3y^3 & x^3y^4 & x^3y^5 = gxy^3 \\
8) & x & xy & xy^2 & xy^3 & x^2y^2 & x^2y^3 & x^3y^2 & x^3y^3 = gx \\
9) & x^2 & x^2y & x^2y^2 & x^2y^3 & x^3y & x^3y^2 & x^4y & x^4y^2 = gx^2 \\
10) & x^3 & x^3y & x^3y^2 & x^3y^3 & x^4y & x^4y^2 & x^5y & x^5y^2 = gx^3 \\
\end{array}
\]

Summations are assumed for all these quantities and \( g \) is written for \( g(xy) \). The equations reduce considerably.

Equation 5 gives us

\[
0_{11} = \frac{1}{12100} \Sigma gxy
\]

1, 3, and 9, combine to give three equations for \( 0_{00}, 0_{02}, \) and \( 0_{20} \), 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

\[
\begin{align*}
1210c_{01} + 21538c_{03} + 12100c &= gy \\
21538c_{01} + 451330c_{03} + 215380c_{21} &= gy^2 \\
12100c_{01} + 215380c_{03} + 215380c_{21} &= gx^2y
\end{align*}
\]
With solutions

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

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

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

also

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

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

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

and from the case \( n = 2 \)

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

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

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

We also have

\[ c_{11} = \frac{1}{12100} \Sigma gxy \]
To six places

$G(xy) =$
\[ + [0.0294554 \Sigma g - 1.05955 \times 10^{-3}(\Sigma g x^2 + \Sigma g y^2)] \\
+ y[6.54859 \times 10^{-3} \Sigma g y - 2.61943 \times 10^{-4} \Sigma g y^3 - 1.05955 \times 10^{-3} \Sigma g x^2 y] \\
+ y^2[1.05955 \times 10^{-4}(\Sigma g y^2 - 10 \Sigma g)] \\
+ y^3[1.47159 \times 10^{-5} \Sigma g y^3 - 2.61943 \times 10^{-4} \Sigma g y] \\
+ xy[8.26445 \times 10^{-5} \Sigma g xy] \\
+ xy^2[1.05955 \times 10^{-5} \Sigma g xy^2 - 1.05955 \times 10^{-4} \Sigma g x] \\
+ x^2y[1.05955 \times 10^{-5} \Sigma g x^2 y - 1.05955 \times 10^{-4} \Sigma g y] \\
+ x[6.54859 \times 10^{-3} \Sigma g x - 2.61943 \times 10^{-4} \Sigma g x^3 - 1.05955 \times 10^{-3} \Sigma g x y^2] \\
+ x^2[1.05955 \times 10^{-4}(\Sigma g x^2 - 10 \Sigma g)] \\
+ x^3[1.47159 \times 10^{-5} \Sigma g x^3 - 2.61943 \times 10^{-4} \Sigma g x] \]
Case $n = 4$

The normal equations are

\[
\begin{array}{cccccccccccccccc}
00 & 01 & 02 & 03 & 04 & 11 & 12 & 13 & 21 & 22 & 31 & 30 & 40 \\
00 & 01 & 02 & 03 & 04 & 11 & 12 & 13 & 21 & 22 & 31 & 30 & 40 \\
1) M y y y y x y x^2 x^3 x^4 y^2 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_y \\
2) y y y y x y x^2 x^3 x^4 y^2 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_y \\
3) y y y y x y x^2 x^3 x^4 y^2 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_y \\
4) y y y y x y x^2 x^3 x^4 y^2 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_y \\
5) y y y y x y x^2 x^3 x^4 y^2 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_y \\
6) x y x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
7) x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
8) x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
9) x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
10) x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
11) x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_{xy} \\
12) x x y x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_x \\
13) y y x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_x \\
14) y x y x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_x \\
15) x y x^2 x^3 x^4 x^5 x^6 x^7 x^8 x^9 x^{10} x^{11} x^{12} x^{13} & = & g_x \\
\end{array}
\]

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 $n = 3$. 

I-11
The new equations to be solved are

\[
\begin{align*}
0_{00} & \quad 0_{02} & \quad 0_{04} & \quad 0_{22} & \quad 0_{20} & \quad 0_{40} \\
M & \quad y^2 & \quad y^4 & \quad x^2 y^2 & \quad x^2 & \quad x^4 & = \Sigma g \\
y^2 & \quad y^4 & \quad y^6 & \quad x^2 y^4 & \quad x^2 y^2 & \quad x^4 y^2 & = \Sigma g y^2 \\
y^4 & \quad y^6 & \quad y^8 & \quad x^2 y^6 & \quad x^2 y^4 & \quad x^4 y^4 & = \Sigma g y^4 \\
x^2 y^2 & \quad x^2 y^4 & \quad x^2 y^6 & \quad x^4 y^4 & \quad x^4 y^2 & \quad x^6 y^2 & = \Sigma g x^2 y^2 \\
x^2 & \quad x^2 y^2 & \quad x^2 y^4 & \quad x^4 y^2 & \quad x^4 & \quad x^6 & = \Sigma g x^2 \\
x^4 & \quad x^4 y^2 & \quad x^4 y^4 & \quad x^6 y^2 & \quad x^6 & \quad x^8 & = \Sigma g x^4
\end{align*}
\]

and

\[
\begin{align*}
c_{11} & \quad c_{13} & \quad c_{31} \\
x^2 y^2 & \quad x^2 y^4 & \quad x^4 y^2 & = \Sigma g x y \\
x^2 y^4 & \quad x^2 y^6 & \quad x^4 y^4 & = \Sigma g x y^3 \\
x^4 y^2 & \quad x^4 y^4 & \quad x^6 y^2 & = \Sigma g x^3 y
\end{align*}
\]

The last set has solutions

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

\[
c_{13} = \frac{1}{679536} [\Sigma g x y^3 - 17.8 \Sigma g x y]
\]

\[
c_{31} = \frac{1}{679536} [\Sigma g x^3 y - 17.8 \Sigma g x y]
\]
The solution of the first set to six places is

\[ a_{00} = 4.53280 \times 10^{-2} \Sigma \epsilon g + 1.58932 \times 10^{-4}(\Sigma \epsilon \gamma^4 + \Sigma \epsilon \gamma^4) \]
\[ + 1.35839 \times 10^{-4}\epsilon \gamma^2 y^2 - 6.399122 \times 10^{-3}(\epsilon \gamma^2 + \epsilon \gamma^2) \]

\[ a_{02} = 1.62141 \times 10^{-3} \epsilon \gamma^2 - 2.81527 \times 10^{-3} \epsilon \gamma \]
\[ - 5.51847 \times 10^{-5} \epsilon \gamma^4 - 1.35839 \times 10^{-5}(\epsilon \gamma^2 y^2 - 10\epsilon \gamma^2) \]

\[ a_{04} = 2.20739 \times 10^{-6} \epsilon \gamma^4 - 5.51847 \times 10^{-5} \epsilon \gamma^2 + 1.58932 \times 10^{-5} \epsilon \gamma \]

\[ a_{22} = 1.35839 \times 10^{-6}(\epsilon \gamma^2 y^2 - 10(\epsilon \gamma^2 + \epsilon \gamma^2) + 100\epsilon \gamma) \]

\[ a_{20} = 1.62141 \times 10^{-3} \epsilon \gamma^2 - 2.81527 \times 10^{-3} \epsilon \gamma \]
\[ - 5.51847 \times 10^{-5} \epsilon \gamma^2 - 1.35839 \times 10^{-5}(\epsilon \gamma^2 y^2 - 10\epsilon \gamma^2) \]

\[ a_{40} = 2.20739 \times 10^{-6} \epsilon \gamma^4 - 5.51847 \times 10^{-5} \epsilon \gamma^2 + 1.58932 \times 10^{-5} \epsilon \gamma \]

To simplify writing \( G(xy) \) we introduce the abbreviations

\[ A = \epsilon \gamma \]
\[ B = \epsilon \gamma x \]
\[ C = \epsilon \gamma x^2 \]
\[ D = \epsilon \gamma x^3 \]
\[ E = \epsilon \gamma x^4 \]
\[ F = \epsilon \gamma xy \]
\[ G = \epsilon \gamma x^2 y \]
\[ H = \epsilon \gamma x^3 y \]
\[ I = \epsilon \gamma xy^2 \]
\[ J = \epsilon \gamma x^2 y^2 \]
\[ K = \epsilon \gamma xy^3 \]
\[ L = \epsilon \gamma y \]
\[ M = \epsilon \gamma^2 \]
\[ N = \epsilon \gamma^3 \]
\[ P = \epsilon \gamma^4 \]
Hence

\[ g(xy) = [4.53280 \times 10^{-2}A + 1.58932 \times 10^{-4}(P + E) \]

\[ + 1.35839 \times 10^{-4}J - 6.39122 \times 10^{-3}(M + C)] \]

\[ + y[6.54859 \times 10^{-3}L - 2.61943 \times 10^{-4}N - 1.05955 \times 10^{-3}G] \]

\[ + y^2[1.62141 \times 10^{-3}M - 2.81527 \times 10^{-3}A - 5.51847 \times 10^{-5}P \]

\[ - 1.35839 \times 10^{-5}(J - 10C)] \]

\[ + y^3[1.47159 \times 10^{-5}N - 2.61943 \times 10^{-4}L] \]

\[ + y^4[2.20739 \times 10^{-6}P - 5.51847 \times 10^{-5}M + 1.58932 \times 10^{-5}A] \]

\[ + xy[1.01516 \times 10^{-3}F - 2.61943 \times 10^{-5}(K + H)] \]

\[ + xy^2[1.05955 \times 10^{-5}I - 1.05955 \times 10^{-4}B] \]

\[ + xy^3[1.47159 \times 10^{-6}K - 2.61943 \times 10^{-5}F] \]

\[ + x^2y[1.05955 \times 10^{-5}G - 1.05955 \times 10^{-4}B] \]

\[ + x^2y^2[1.35839 \times 10^{-6}(J - 10(C + M) + 100A)] \]

\[ + x^3y[1.47159 \times 10^{-6}H - 2.61943 \times 10^{-5}F] \]

\[ + x[6.54859 \times 10^{-3}B - 2.61943 \times 10^{-4}D - 1.05955 \times 10^{-3}I] \]

\[ + x^2[1.62141 \times 10^{-3}C - 2.81527 \times 10^{-3}A - 5.51847 \times 10^{-5}C \]

\[ - 1.35839 \times 10^{-5}(J - 10M)] \]

\[ + x^3[1.47159 \times 10^{-5}D - 2.61943 \times 10^{-4}B] \]

\[ + x^4[2.20739 \times 10^{-6}E - 5.51847 \times 10^{-5}C - 1.58932 \times 10^{-5}A] \]
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 \( (c_{00}, c_{01}, c_{02}, c_{04}, c_{10}, c_{11}, c_{13}, c_{31}) \) 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 $\Sigma g_{ij}$, 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 $a_i$ and $b_i$ 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 $\Sigma g$, $\Sigma g_x$, $\Sigma g_x^2$, $\Sigma g_x^3$, $\Sigma g_y$, $\Sigma g_y^2$, $\Sigma g_y^3$, from the relations

\[
\begin{align*}
\Sigma g &= \Sigma a_i \\
\Sigma g_x &= \Sigma a_i (i) \\
\Sigma g_x^2 &= \Sigma a_i (i)^2 \\
\Sigma g_x^3 &= \Sigma a_i (i)^3 \\
\Sigma g_y &= \Sigma b_i (1) \\
\Sigma g_y^2 &= \Sigma b_i (1)^2 \\
\Sigma g_y^3 &= \Sigma b_i (1)^3
\end{align*}
\]

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

For the remaining summations $\Sigma gx$, $\Sigma gx^2y$, $\Sigma gxy^2$ 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.
Fig. 1.3

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.

**Example**

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.
RESIDUAL GRAVITY WITH SECOND ORDER POLYNOMIAL

$$G(\omega) = 4187 - 8322\omega - 391\omega^2 - 0.557\omega^3$$

+ 2.904$\omega^4$ - 395$\omega^5$

Fig. 1.6
RESIDUAL GRAVITY WITH THIRD ORDER POLYNOMIAL

Fig. 1.7
RESIDUAL GRAVITY WITH
FOURTH ORDER POLYNOMIAL

\[ G(mg) = 92.09 - 88.2y - .251z^2 - .0557z^3 - .0075z^4 - .395w + .0203w^2 + .0005w^3 + .001w^4 + .0001w^5 + .0097x^2 + 1.75x - .584x^2 + .0339x^3 + .0044x^4 \]

Fig. 1.8
RESIDUAL GRAVITY
COMPANY DATA

Fig. 1.9
CONTOURS OF $G(x,y)$

2nd ORDER POLYNOMIAL

3rd ORDER POLYNOMIAL

4th ORDER POLYNOMIAL

Fig. 1.10
REGIONAL CORRECTION MINUS TOPOGRAPHIC CORRECTION

COMPANY DATA

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:

\[
\begin{align*}
\Sigma x^2 & \quad \Sigma x^4 \quad \Sigma x^6 \quad \Sigma y^4 \quad \Sigma y^6 \quad \Sigma y^8 \\
\Sigma x^2y^2 & \quad \Sigma x^4y^4 \quad \Sigma x^2y^6 \quad \Sigma x^4y^2 \quad \Sigma x^6y^2 \quad \Sigma x^0y^0
\end{align*}
\]

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 \( \Sigma x^k \) over the grid is equal to the \( \Sigma x^k \) 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

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

Likewise

\[
\Sigma y^k = 2(2N + 1)(\sum_{i=1}^{M} i^k)
\]
For the cross terms $\Sigma x^k y^l$ we have

$$\Sigma \frac{x^k y^l}{\text{grid}} = \Sigma_{j=-M}^{N} \Sigma_{i=-N}^{M} i^k j^l$$

$$= \left[ \Sigma i^k \right] \left[ \Sigma j^l \right]$$

$$= \left[ \Sigma_{i=-N}^{N} i^k \right] \left[ \Sigma_{j=-M}^{M} j^l \right]$$

$$\Sigma \frac{x^k y^l}{\text{grid}} = 4 \left( \Sigma i^k \right) \left( \Sigma j^l \right)$$

Hence the sums $1.3$ are easily derivable from equations $1.4$ and $1.5$ if we tabulate the quantities $\Sigma i^k$. Table I gives values of $i^k$ from which the sums $\Sigma i^k$ 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>
<thead>
<tr>
<th>$k$</th>
<th>$1^k$</th>
<th>$2^k$</th>
<th>$4^k$</th>
<th>$6^k$</th>
<th>$8^k$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>4</td>
<td>16</td>
<td>64</td>
<td>256</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>9</td>
<td>81</td>
<td>729</td>
<td>6561</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>16</td>
<td>256</td>
<td>4096</td>
<td>65536</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>625</td>
<td>15625</td>
<td>390625</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>1296</td>
<td>46656</td>
<td>1679616</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>49</td>
<td>2401</td>
<td>117649</td>
<td>5764801</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>64</td>
<td>4096</td>
<td>262144</td>
<td>16777216</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>81</td>
<td>6561</td>
<td>531441</td>
<td>43046721</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>100</td>
<td>10000</td>
<td>1000000</td>
<td>1000000000</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>121</td>
<td>14641</td>
<td>1771561</td>
<td>214358881</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>144</td>
<td>20736</td>
<td>2985984</td>
<td>429981696</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>169</td>
<td>28561</td>
<td>4826809</td>
<td>815730721</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>196</td>
<td>38416</td>
<td>7529536</td>
<td>1475789056</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>225</td>
<td>50625</td>
<td>11390625</td>
<td>2562890625</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>256</td>
<td>65536</td>
<td>16777216</td>
<td>4294967296</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>289</td>
<td>83521</td>
<td>24137569</td>
<td>6975757441</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>324</td>
<td>104976</td>
<td>34012224</td>
<td>11019960576</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>361</td>
<td>130321</td>
<td>47045881</td>
<td>16983563041</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>400</td>
<td>160000</td>
<td>64000000</td>
<td>25600000000</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>441</td>
<td>194481</td>
<td>85766121</td>
<td>37822859361</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>484</td>
<td>234256</td>
<td>113379904</td>
<td>54875873536</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>529</td>
<td>279841</td>
<td>148035889</td>
<td>78310985281</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>576</td>
<td>331776</td>
<td>191102976</td>
<td>110075314176</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>625</td>
<td>390625</td>
<td>244140625</td>
<td>152587890625</td>
<td></td>
</tr>
<tr>
<td>$L$</td>
<td>$k=2$</td>
<td>$k=4$</td>
<td>$k=6$</td>
<td>$k=8$</td>
<td></td>
</tr>
<tr>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>17</td>
<td>65</td>
<td>257</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14</td>
<td>98</td>
<td>794</td>
<td>6818</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>30</td>
<td>354</td>
<td>4890</td>
<td>72354</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>55</td>
<td>979</td>
<td>20515</td>
<td>462979</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>91</td>
<td>2275</td>
<td>67171</td>
<td>2142595</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>140</td>
<td>4676</td>
<td>184820</td>
<td>7907396</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>204</td>
<td>8772</td>
<td>446964</td>
<td>24684612</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>285</td>
<td>15333</td>
<td>978405</td>
<td>67731333</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>385</td>
<td>25333</td>
<td>1978405</td>
<td>167731333</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>506</td>
<td>39974</td>
<td>3749966</td>
<td>382090214</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>650</td>
<td>60710</td>
<td>6735950</td>
<td>812071910</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>819</td>
<td>89271</td>
<td>11562759</td>
<td>1627802631</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>1015</td>
<td>127687</td>
<td>19092295</td>
<td>3103591687</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>1240</td>
<td>178312</td>
<td>30482920</td>
<td>5666482312</td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>1496</td>
<td>243848</td>
<td>47260136</td>
<td>9961449608</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>1785</td>
<td>327369</td>
<td>71397705</td>
<td>16937207049</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2109</td>
<td>432345</td>
<td>105409929</td>
<td>27957167625</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>2470</td>
<td>562666</td>
<td>152455810</td>
<td>44940730666</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2870</td>
<td>722666</td>
<td>216455810</td>
<td>70540730666</td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>3311</td>
<td>917147</td>
<td>302221931</td>
<td>108363590027</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>3795</td>
<td>1151403</td>
<td>415601835</td>
<td>163239463563</td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>4324</td>
<td>1431244</td>
<td>563637724</td>
<td>241550448844</td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>4900</td>
<td>1763020</td>
<td>754740700</td>
<td>351625763020</td>
<td></td>
</tr>
<tr>
<td>25</td>
<td>5525</td>
<td>2153645</td>
<td>998881325</td>
<td>504213653645</td>
<td></td>
</tr>
<tr>
<td>N=5</td>
<td>M=5</td>
<td>N=5</td>
<td>M=10</td>
<td>N=10</td>
<td>M=10</td>
</tr>
<tr>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>$\Sigma x^0 y^0$</td>
<td>121</td>
<td>231</td>
<td>441</td>
<td>961</td>
<td>1681</td>
</tr>
<tr>
<td>$\Sigma x^2$</td>
<td>1210</td>
<td>2310</td>
<td>16170</td>
<td>76880</td>
<td>235340</td>
</tr>
<tr>
<td>$\Sigma x^4$</td>
<td>21538</td>
<td>41118</td>
<td>1063986</td>
<td>11055344</td>
<td>59258612</td>
</tr>
<tr>
<td>$\Sigma x^6$</td>
<td>451330</td>
<td>861630</td>
<td>83093010</td>
<td>1889941040</td>
<td>17749376420</td>
</tr>
<tr>
<td>$\Sigma x^8$</td>
<td>10185538</td>
<td>19445118</td>
<td>7044715986</td>
<td>351321903344</td>
<td>5784339914612</td>
</tr>
<tr>
<td>$\Sigma y^2$</td>
<td>1210</td>
<td>8470</td>
<td>16170</td>
<td>76880</td>
<td>235340</td>
</tr>
<tr>
<td>$\Sigma y^4$</td>
<td>21538</td>
<td>557326</td>
<td>1063986</td>
<td>11055344</td>
<td>59258612</td>
</tr>
<tr>
<td>$\Sigma y^6$</td>
<td>451330</td>
<td>43524910</td>
<td>83093010</td>
<td>1889941040</td>
<td>17749376420</td>
</tr>
<tr>
<td>$\Sigma y^8$</td>
<td>10185538</td>
<td>3690089326</td>
<td>7044715986</td>
<td>351321903344</td>
<td>5784339914612</td>
</tr>
<tr>
<td>$\Sigma x^2 y^2$</td>
<td>12100</td>
<td>84700</td>
<td>592900</td>
<td>6150400</td>
<td>32947600</td>
</tr>
<tr>
<td>$\Sigma x^2 y^4$</td>
<td>215380</td>
<td>5573260</td>
<td>39012820</td>
<td>884427520</td>
<td>8296205680</td>
</tr>
<tr>
<td>$\Sigma x^2 y^6$</td>
<td>4513300</td>
<td>435249100</td>
<td>3046743700</td>
<td>151195283200</td>
<td>2484912698800</td>
</tr>
<tr>
<td>$\Sigma x^4 y^2$</td>
<td>215380</td>
<td>1507660</td>
<td>39012820</td>
<td>884427520</td>
<td>8296205680</td>
</tr>
<tr>
<td>$\Sigma x^4 y^4$</td>
<td>4513300</td>
<td>31593100</td>
<td>3046743700</td>
<td>151195283200</td>
<td>2484912698800</td>
</tr>
<tr>
<td>$\Sigma x^4 y^6$</td>
<td>3833764</td>
<td>99204028</td>
<td>2567043556</td>
<td>127180677376</td>
<td>2088984590224</td>
</tr>
</tbody>
</table>

Table III: $\Sigma_{k=0, 2, 4, 6, 8} \Sigma_{l=0, 2, \ldots, (8-k)}$ for various grids measuring $2N \times 2M$. 
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 instrumental. 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 basically 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 Refs. 5 and 6. The linear operator permits a measure of the change in dynamics as we proceed down a seismic record. As these dynamics 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:

\[
\hat{x}_{i+2} = (1 - a - b)\bar{x} + 2(l-u)\bar{x} + x_i + 2u(x_{i+1} - \bar{x})
\]

where

\[
\begin{align*}
a &= 2 \cos 2\pi hf = 2u \\
b &= 1 \\
h &= \text{time between observation} \\
c &= (1-a-b)\bar{x} = (1-u)\bar{x} \\
\bar{x} &= \text{mean of series} \\
f &= \text{frequency of cosine wave} \\
\hat{x}_{i+2} &= \text{predicted value of } x_{i+2}
\end{align*}
\]

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

\[
E_{i+2} = x_{i+2} - [2(l-u)\bar{x} + 2ux_{i+1} - x_i]
\]

\[
= x_{i+2} - 2(l-u)\bar{x} - 2ux_{i+1} + x_i
\]

\[
= (x_{i+2} - \bar{x}) + (x_i - \bar{x}) - 2u(x_{i+1} - \bar{x})
\]

For simplicity let us deal with a series \( X_1 \) measured around its equilibrium mean \( \bar{x} \), i.e. \( X_1 = x_i - \bar{x} \) then 2.2 becomes

\[
E_{i+2} = X_{i+2} + X_1 - 2uX_{i+1}
\]

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

\( \dagger \) Ref. 6
\[\sum_{i} E_{i+2}^2 = \sum (X_{i+2}^2 + X_i^2 + 4u^2 X_{i+1}^2 + 2X_{i+2} X_i - 4uX_{i+2} X_i + 1) \quad 2.4\]

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 \(\sum X_i^2 = 1\) then

\[\sum_{i} E_{i+2}^2 = R_0 + R_0 + 4u^2 R_0 + 2R_2 - 8uR_1 \quad 2.5\]

where \(R_1 = \text{1-th lag auto-correlation and } R_0 = 1\)

or

\[\sum_{i} E_{i+2}^2 = 2(1 + R_2 - 4uR_1 + 2u^2) \quad 2.6\]

This expression has a minimum value when

\[u = \frac{-(-4R_1)}{2 \times 2} = R_1 \quad 2.7\]

or

\[\cos 2\pi hf = R_1 \quad 2.8\]

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

where \(\cos^{-1}\) 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
\[
\text{Min } \sum E_{1+2}^2 = 2(1 + R_2^2 - 2R_1^2)
\]  

If we want a perfect least squares fit we have
\[
R_1 = \pm \sqrt{\frac{1}{2} (1 + R_2^2)}
\]

with the restriction \(2\pi hf = R_1\)

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

The Geometry of Cosine Operators

A. Error Sum as Function of \(W\)

Consider the sum of squared errors as a function of \(u\). We have
\[
\sum E_{1+2}^2 = 2(1 + R_2 - 4uR_1 + 2u^2)
\]

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 \(-1 < R_1 < +1\), \(\sum E^2\) 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
We require \( \Sigma E^2 \) to be non-negative. This means the discriminant of 2.9 must be \( \leq 0 \) or

\[
16 R_1^2 - 8(1+R_2) \leq 0
\]

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

\[
E_{i+2} = X_{i+2} + X_i - 2uX_{i+1}
\]

If we fix attention on a single individual error \((i \text{ constant})\) and let \( u \) vary we see that \( E_{i+2} \) is sinusoidal since \( u = \cos 2\pi hf \). Thus \( E_{i+2} \) varies sinusoidally about a mean given by the sum of the \( i^{th} \) and the \((i-2)^{th}\) value of the series, with an amplitude of twice the \((i-1)^{st}\) 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 $\Sigma E^2$ is parabolic, determining 3 values of $\Sigma E^2$ 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 $\sum E^2$.

**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 ops. 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 $\cot^{-1}(\mu)$

$\sum (R-x)^2$

Fig. 2.5
EXAMPLES OF THE SINUSOIDAL
DISTRIBUTION OF INDIVIDUAL
ERRORS FOR COSINE OPERATORS
AS FUNCTIONS OF $\mu$
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 $h = \frac{1}{2}$ 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.

\[ E_{1+2} = X_{1+2} + X_1 - 2\mu X_{1+1} \]

To get the spectrum of this function we first find the auto-correlation \( R_\tau \).

\[
R_\tau = \sum (X_{1+2} + X_1 - 2\mu X_{1+1})(X_{1+2-\tau} + X_{1-\tau} - 2\mu X_{1+1-\tau})
\]

\[
R_\tau = \sum X_{1+2}X_{1+2-\tau} + \sum X_{1+2}X_{1-\tau} + \sum X_1X_{1+2-\tau}
\]

\[
+ \sum X_1X_{1-\tau} - 2\mu (\sum X_1X_{1+2-\tau} + \sum X_1X_{1-\tau})
\]

\[
+ \sum X_{1+1-\tau}X_{1+2} + \sum X_{1+1-\tau}X_1 + 4\mu^2 \sum X_{1+1}X_{1+1-\tau}
\]

If the \( X \) series is properly normalized we may write this in terms of the correlations \( r_\tau \) of the \( X \) series.

\[
R_\tau = r_\tau + r_{\tau+2} + r_{\tau-2} + 4\mu^2 r_\tau
\]

\[
+ r_\tau - 2\mu (r_{\tau-1} + r_{\tau+1} + r_{\tau+1} + r_{\tau-1})
\]

\( \dagger \) Ref. 7
\[ \begin{align*}
R_T &= r_{T-2} + r_{T-1}(-4u) + r_T(2+4u^2) \\
&\quad + r_{T+1}(-4u) + r_{T+2} \\
\text{Since the series is taken to be autoregressive,} \\
r_T &= \cos2\pi f_0 T e^{-a T} = \cos at e^{-bt} \quad \forall \ T > 0
\end{align*} \]

Substituting equation 2.14 into 2.13 we have

\[ \begin{align*}
R_T &= \cos[a(T-2)]e^{-b(T-2)} \\
&\quad - 4ucos[a(T-1)]e^{-b(T-1)} \\
&\quad + (2+4u^2)\cos at e^{-bt} \\
&\quad - 4ucos[a(T+1)]e^{-b(T+1)} \\
&\quad + \cos[a(T+2)]e^{-b(T+2)}
\end{align*} \]

Using trigonometric identities

\[ \begin{align*}
R_T &= e^{2b} [\cos at \cos 2a + \sin at \sin 2a]e^{-bt} \\
&\quad - 4ue^{b} [\cos at \cos a + \sin at \sin a]e^{-bt} \\
&\quad + (2+4u^2)(\cos at)e^{-bt} \\
&\quad - 4ue^{-b} [\cos at \cos a - \sin at \sin a]e^{-bt} \\
&\quad + e^{-2b} [\cos at \cos 2a - \sin at \sin 2a]e^{-bt}
\end{align*} \]

or

\[ \begin{align*}
R_T &= \cos at e^{-bt} [e^{2b} \cos 2a - 4ue^{b} \cos a + 2 + 4u^2 \\
&\quad - 4ue^{-b} \cos a + e^{-2b} \cos 2a]
\end{align*} \]

† Ref: 5

II-12
\[+ \sin \alpha \e^{-bT} [e^{2b}\sin 2a - 4ue^b \sin a + 4ue^{-b} \sin a - e^{-2b} \sin 2a]\]

Hence

\[R_T = A \cos \alpha \e^{-bT} + B \sin \alpha \e^{-bT}\]

\[= \e^{-bT} (A \cos \alpha + B \sin \alpha) \quad 2.15\]

where

\[A = \cos 2\alpha (e^{2b} + e^{-2b}) - 4\cos \alpha (e^b + e^{-b}) + 2 + 4u^2\]

\[B = \sin 2\alpha (e^{2b} - e^{-2b}) - 4\sin \alpha (e^b - e^{-b}) \quad 2.16\]

\[R_T = \e^{-bT} (A^2 + B^2)^{1/2} \left[ \frac{A}{(A^2 + B^2)^{1/2}} \cos \alpha + \frac{B}{(A^2 + B^2)^{1/2}} \sin \alpha \right]\]

\[= \e^{-bT} (A^2 + B^2)^{1/2} (\cos \alpha \cos \beta + \sin \alpha \sin \beta)\]

where \(\beta = \tan^{-1} \frac{B}{A}\)

Thus

\[R_T = \e^{-bT} (A^2 + B^2)^{1/2} \cos (\alpha + \beta) \quad 2.17\]

\(R_T\) is now in a form similar to the \(r_T\) 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 \(\beta\).

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.

II-13
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:

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

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

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

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

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

or substituting:

$$R_\tau = e^{\frac{-\sigma^2 (\tau-2)^2}{2}} \cos[a(\tau-2)]$$

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

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

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

$$+ e^{\frac{-\sigma^2 (\tau+2)^2}{2}} \cos[a(\tau+2)]$$

† Ref. 5

II-14
This can be reduced, as before, to the form

\[ R(\tau) = e^{-\frac{\tau^2}{2}} \left[ A(\tau) \cos \alpha \tau + B(\tau) \sin \alpha \tau \right] \tag{2.21} \]

where

\[ A(\tau) = \cos 2\alpha \left( e^{-\frac{\tau^2}{2}} + e^{-2\tau^2 + 1} \right) - 4\alpha \cos \alpha \left( e^{-\frac{\tau^2}{2}} + e^{-2\tau^2 + 1} \right) + (2 + 4\alpha^2) \tag{2.22} \]

\[ B(\tau) = \sin 2\alpha \left( e^{-\frac{\tau^2}{2}} + e^{-2\tau^2 + 1} \right) - 4\alpha \sin \alpha \left( e^{-\frac{\tau^2}{2}} + e^{-2\tau^2 + 1} \right) \]

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

\[ \phi(\omega) = 2 \int_0^\infty R(\tau) \cos \omega \tau \, d\tau \tag{2.23} \]

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 $\sigma = 224\text{cps}$, $\nu = 50\text{ cps}$ (normalized).

----- 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 \( \sigma = 224 \text{ ops} \) \( \Delta = 50 \text{ ops} \) (normalized).

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

\[ h = 2.5 \times 10^{-3} \text{ sec.} \]
Fig. 2.10
FILTER CHARACTERISTICS
FOR 50 CYCLE COSINE OPERATOR
WITH $h = 2.5\ ms$. 

°/o Power removed / 100

\[
\begin{align*}
&1.0 \\
&0.9 \\
&0.8 \\
&0.7 \\
&0.6 \\
&0.5 \\
&0.4 \\
&0.3 \\
&0.2 \\
&0.1 \\
\end{align*}
\]
\[
\begin{align*}
&10 20 30 40 50 60 70 80 90 100 \\
&\text{CPS} \rightarrow
\end{align*}
\]
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

\[
\hat{x}_{i+k} = c + a_0 x_i + a_1 x_{i-1} + \ldots + a_M x_{i-M} \\
+ b_0 y_i + \ldots + b_M y_{i-M} \\
+ c_0 z_i + \ldots + c_M z_{i-M} \\
+ d_0 u_i + \ldots + d_M u_{i-M}
\] 2.24

\[I = \sum (x_{i+k} - \hat{x}_{i+k})^2\] is a minimum. 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 \( a = \overline{x} \) (mean of \( x_i \) 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 \( x_i \) series were truly random and had no predictability. With these values of the constants \( I = I(\overline{x}, 0, 0, \ldots) \) becomes the sample variance about the sample mean.

The computational procedure is:

1. Find \( I(\overline{x}, 0, 0, \ldots) \)
2. Find \( I(\overline{x} + \Delta a, 0, 0, \ldots) \)
3. If \( 2 < 1 \), continue adding \( \Delta a \) until
   \[ I(\overline{x} + n \Delta a, 0, 0, \ldots) > I(\overline{x} + (n-1) \Delta a, 0, 0, \ldots) \]
   If \( 2 > 1 \), subtract \( \Delta a \) and continue to subtract until
   \[ I(\overline{x} - n \Delta a, 0, 0, \ldots) > I(\overline{x} - (n-1) \Delta a, 0, 0, \ldots) \]
4. Using \( i + (n-1) \Delta a, 0, 0, \ldots \), as the new starting point, find \( I(i + (n-1) \Delta a, \Delta a_0, 0, 0, \ldots) \)
   and repeat the steps under 3.
5. Work successively in this fashion with each of the variables \( a, a_0, a_1 \ldots a_M \).
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 elliptical. 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
through 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.
\[ I = \sum (R_{i+2} - R_{i+1})^2 \]

Where

\[ R_{i+2} = a + a_1 y_i + a_2 y_{i+1} + a_3 y_{i+2} + a_4 y_{i+3} + d_1 u_i + d_2 u_{i+1} + d_3 u_{i+2} + d_4 u_{i+3} \]

Initial Value of I

\[ \hat{R}_{1,1} = \frac{V}{2} \]
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 0. 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

\[
\begin{array}{c}
I \\
II
\end{array}
\]

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 0. 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
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 \(OE - \frac{1}{4} AD\). 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 sufficient 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 \( V_A \neq V_B \), 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

\[ + \text{Ref. 8} \]

III-2
\[ F_N = A_N + B_N \]  
\[ G_N = A_N + B_{N-j} \]  
\[ H_N = A_N + B_{N-2j} \]

where the time lag between traces is approximated by \( j \) units so that \[ \frac{J}{L} = \frac{-d}{V_A} + \frac{d}{V_B} \]

or
\[ V_B = \frac{JdV_A}{jV_A + Jd} \]

This is illustrated in Fig. 3.1.

From equations 3.1, 3.2, and 3.3 we can get
\[ A_N - A_{N-j} = G_N - F_{N-j} \]  
\[ A_N - A_{N-2j} = H_N - F_{N-2j} \]

3.5 and 3.6 are recursion formulas giving \( A_{N-kj} \) and \( A_{N-2kj} \) respectively \( (k = 1, 2, \ldots) \) once \( A_N \ldots A_{N-j+1} \) are known. Now if \( j \) has its correct value then it is easy to show that regardless of how we choose the initial \( A \)'s 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 \( A_N, A_{N-1}, \ldots A_{N-j+1} \)

use equations 3.5 and 3.6 to calculate the two series (to
\(F_N = A_N + B_N\)
\(G_N = A_N + B_{N-J}\)
\(H_N = A_N + B_{N-2J}\)

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.

\[ A_N - A_{N-j} = G_N - F_{N-j} \]  

Equation 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

$$3A_{N+K} = A_N + A_{N+J} + A_{N+K-J} + G_{N+K}$$

$$+ 2H_{N+K} - F_N - F_{N+K-J} - G_{N+J}$$

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

**Tests of the Method**

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

III-6
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 entire 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
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

III-8
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

--- Figure 3.6 ---

Amplitude vs. Reflection

Time (Seconds)
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 amplitudes 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:
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.
PHASE TEST
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 rate of change 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 lth trace (l = 1, 2, ...T) we have chosen N_l operators. For the kth operator on this trace (k = 1, ...N_l) there is an associated error time series which we define as e_{1k}(l). Then, for example, we may construct a single error time series $\varepsilon_i^l$ to be associated with the lth trace by the expression

$$\varepsilon_i^l = \sum_{k=1}^{N_l} [e_{1k}(l)]^2$$

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 $\delta_i^x(\alpha)$ with an arbitrary lag or lead $\alpha$.

III-13
\[ \delta_{(a)} = \sum_{t=1}^{T} \epsilon_{1-a^t} \quad a = 0, \pm 1, \pm 2, \ldots 3.8 \]

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 \( a \) 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.

III-15
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


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

\[
g(xy) = c_{00} + c_{10}x + c_{20}x^2 + \ldots + c_{n0}x^n
\]

\[
+ c_{01}y + c_{11}xy + \ldots + c_{(n-1),1}x^{n-1}y
\]

\[
= \vdots \vdots \vdots
\]

\[
+ c_{0,(n-1)y^{n-1}} + c_{1,(n-1)}xy^{n-1}
\]

\[
+ c_{0n}y^n
\]

where the \( c_{ik} \)'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 +N greatest value of y

The coefficients $C_{4k}$ of the polynomial must be scale factored in a special way because they decrease in magnitude rapidly as $x+k$ increases. The scale factor is $10^{(x+k)-2}$, 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 $c_{00} \times 10^{-2}$ 461 $c_{12} \times 10$
(Octal) 444 $c_{10} \times 10^{-2}$ 462 $c_{22} \times 10^{2}$
        445 $c_{20} \times 1$ 463 $c_{32} \times 10^{3}$
        446 $c_{30} \times 10$ 464 $c_{42} \times 10^{4}$
        447 $c_{40} \times 10^{2}$ 465 $c_{03} \times 10$
        450 $c_{50} \times 10^{3}$ 466 $c_{13} \times 10^{2}$
        451 $c_{60} \times 10^{4}$ 467 $c_{23} \times 10^{3}$
        452 $c_{01} \times 10^{-1}$ 470 $c_{33} \times 10^{4}$
        453 $c_{11} \times 1$ 471 $c_{04} \times 10^{2}$
        454 $c_{21} \times 10$ 472 $c_{14} \times 10^{3}$
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:

<table>
<thead>
<tr>
<th>Register</th>
<th>Equation</th>
</tr>
</thead>
<tbody>
<tr>
<td>540</td>
<td>( g(-N,M) \times 10^{-2} )</td>
</tr>
<tr>
<td>541</td>
<td>( g(-N+1,M) \times 10^{-2} )</td>
</tr>
<tr>
<td>542</td>
<td>( g(-N+2,M) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( \ldots )</td>
</tr>
<tr>
<td></td>
<td>( g(N,M) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( g(-N,M-1) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( g(-N+1,M-1) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( \ldots )</td>
</tr>
<tr>
<td></td>
<td>( g(N,M-1) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( g(-N,M-2) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( \ldots )</td>
</tr>
<tr>
<td></td>
<td>( g(-N,-M) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( g(-N+1,-M) \times 10^{-2} )</td>
</tr>
<tr>
<td></td>
<td>( \ldots )</td>
</tr>
</tbody>
</table>
Now suppose the information \( n, N, M \), and the 
the \( C_{jk} \)'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 l
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.

A Technical Feature in Polynomial I

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 \( I+k \) (the \( C_{ik} \)'s) by numbers rapidly increasing in magnitude with \( I+k \) \((x^I y^k)\) 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 \( C_{ik} x^I y^k \) to be scale factored finally by \( 10^{-2} \). To preserve accuracy during the computation of the product we do the following:

1. Form \( x^I 2^{-15} \) and \( y^k 2^{-15} \) and then scale factor to \( x^I 2^{-15+\omega} \) and \( y^k 2^{-15+\beta} \) by use of the scale factor order.

2. Form \( C_{ik} 10(I+k)2x^I y^k 2^{-15+\alpha k} 2^{-15+\beta} \)

\[
= C_{ik} x^I y^k 2^{-30+\alpha+\beta} 10(I+k)2 \tag{1}
\]

To get this product to the form \( C_{ik} x^I y^k 10^{-2} \) we must multiply by \( 10^{-((I+k)230-(\alpha+\beta))} \)

It appears that we merely need to store the negative powers of 10, multiply the expression (1) by \( 10^{-((I+k)230-(\alpha+\beta))} \) and then shift left \( 30-(\alpha+\beta) \). However the negative powers of 10 cannot be stored with any accuracy for high \( I+k \) so we write \( 10^{-((I+k)230-(\alpha+\beta))} \) in the form
\[-(\lambda + k) \log_{10} 30 - (\alpha + \beta) \]
\[= \frac{1}{2} - 3.32193(\lambda + k) + 30 - (\alpha + \beta) \]
\[= \frac{1}{2} - 3.32193(\lambda + k) \cdot [2 - 3(\lambda + k) + 30 - (\alpha + \beta)] \]
\[= (0.800)^{\lambda + k} \cdot [2 - 3(\lambda + k) + 30 - (\alpha + \beta)] \]

We can store the powers of (0.800) with ample accuracy. Thus we multiply by the appropriate power of (0.800) and follow this by a shift left or right according to the exponent of 2. (The zeroth power of (0.800) is put in as +0.9999.)
\[ m = 4 \]

\[
\begin{array}{cccccccccccccccc}
+0385 & -0302 & -0008 & -0147 & -0094 & -0136 & +0082 & +0153 & +0209 & +0173 & -0146 \\
+0163 & +0004 & -0325 & +0062 & -0049 & -0193 & -0007 & -0045 & +0031 & +0001 & -0120 \\
+0086 & -0132 & -0082 & +0133 & +0553 & +0301 & +0056 & +0116 & +0063 & +0069 & -0078 \\
-0293 & -0151 & -0025 & +0141 & +0050 & +0071 & +0217 & +0250 & -0145 & +0111 & -0074 \\
-0145 & -0164 & -0079 & +0376 & +0563 & +0179 & -0136 & -0208 & -0220 & +0084 & +0149 \\
+0099 & -0028 & +0070 & +0071 & +0138 & +0175 & +0222 & +0214 & +0224 & +0037 & +0165 \\
+0249 & +0225 & +0093 & +0149 & +0151 & -0045 & -0231 & -0123 & -0004 & +0030 & +0098 \\
+0095 & +0060 & +0004 & +0097 & +0264 & +0241 & +0070 & +0011 & +0173 & +0104 & +0044 \\
-0157 & +0417 & -0432 & +0112 & +0106 & +0097 & +0220 & +0224 & +0350 & +0125 & -0583 \\
\end{array}
\]

\[ m = 3 \]

\[
\begin{array}{cccccccccccccccc}
+0372 & -0293 & +0001 & -0172 & -0181 & +0300 & +0154 & +0115 & +0007 & +0146 & +0223 \\
+0193 & +0055 & -0211 & +0179 & +0035 & +0166 & -0051 & -0140 & -0058 & +0021 & +0175 \\
+0007 & +0030 & +0070 & +0071 & +0138 & +0175 & +0222 & +0214 & +0224 & +0037 & +0165 \\
-0029 & -0093 & +0003 & +0350 & +0019 & +0036 & +0027 & -0179 & -0108 & +0032 & +0124 \\
-0063 & -0292 & -0050 & +0138 & +0100 & +0063 & +0196 & +0253 & +0070 & +0001 & +0034 \\
-0249 & -0296 & +0170 & +0355 & +0065 & +0059 & +0108 & -0169 & +0248 & -0018 & -0022 \\
+0012 & +0112 & -0008 & +0047 & +0181 & +0270 & -0108 & +0030 & +0117 & +0056 & +0070 \\
+0040 & +0057 & -0143 & +0122 & +0046 & +0100 & +0012 & +0062 & +0075 & +0070 & +0074 \\
+0390 & +0041 & -0095 & +0193 & -0309 & +0115 & +0014 & +0106 & +0288 & +0186 & +0078 \\
+0212 & +0390 & +0245 & +0302 & +0131 & +0067 & +0133 & +0200 & +0358 & +0134 & -0595 \\
\end{array}
\]

\[ m = 2 \]

\[
\begin{array}{cccccccccccccccc}
+0105 & -0402 & +0035 & -0189 & -0217 & -0363 & -0242 & -0189 & -0009 & +0297 & +0632 \\
+0044 & +0079 & -0189 & +0231 & +0043 & -0203 & -0138 & -0217 & -0015 & +0075 & +0464 \\
-0000 & -0019 & -0136 & +0421 & +0731 & +0369 & -0115 & -0273 & +0104 & +0041 & +0264 \\
-0075 & +0001 & +0099 & +0478 & +0363 & -0033 & -0224 & -0288 & -0224 & -0134 & +0082 \\
-0333 & -0164 & +0181 & +0504 & +0104 & +0030 & -0065 & +0334 & +0277 & -0192 & -0030 \\
-0372 & +0158 & -0130 & +0298 & +0192 & +0063 & -0283 & +0413 & -0111 & -0131 & +0075 \\
-0293 & -0185 & -0001 & +0510 & +0699 & +0264 & -0192 & -0323 & +0426 & -0152 & -0002 \\
+0003 & -0043 & +0134 & +0190 & +0273 & +0283 & -0179 & +0165 & +0274 & -0056 & +0088 \\
+0218 & +0015 & +0133 & +0061 & -0033 & +0070 & -0148 & +0089 & -0054 & +0007 & +0147 \\
+0102 & +0106 & -0039 & -0093 & +0022 & -0073 & +0002 & +0054 & +0234 & +0191 & +0225 \\
-0195 & +0239 & -0698 & -0309 & -0042 & +0000 & +0169 & +0217 & +0391 & +0243 & -0328 \\
\end{array}
\]
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 \( x_i(\text{y}_i, z_i, \text{or u}_i) \) with a linear operator of the general form

\[
x_{i+k}(\text{y}_{i+k}, \text{z}_{i+k} \text{or u}_{i+k}) = a + a_{0}x_{i} + a_{1}x_{i-1} + \cdots + a_{M}x_{i-M} \\
+ b_{0}y_{i} + \cdots + b_{M}y_{i-M} \\
+ c_{0}z_{i} + \cdots + c_{M}z_{i-M} \\
+ d_{0}u_{i} + \cdots + d_{M}u_{i-M}
\]

where the prediction distance \( k \) and the number of lags \( M \) are arbitrary but have the restrictions that \( M \geq 7 \) and \( M+k \leq 19 \). The four series which this program handles contain 500 members each so that \( i \) ranges from 0 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

\[
\sum_{i=j-5}^{i=j+4} (x_i - x_j)^2 \quad j = 25, 35, 45, 55, \ldots, 495
\]

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 \((x_i - x_1)\) 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 form $x_1 = \bar{x}$ (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.
<table>
<thead>
<tr>
<th>12.4 Second Half</th>
<th>12.4 Second Half</th>
</tr>
</thead>
<tbody>
<tr>
<td>$+1300$</td>
<td>$+1400$</td>
</tr>
<tr>
<td>$+0027$</td>
<td>$-0023$</td>
</tr>
<tr>
<td>$+0036$</td>
<td>$+0016$</td>
</tr>
<tr>
<td>$-0009$</td>
<td>$+0030$</td>
</tr>
<tr>
<td>$+0027$</td>
<td>$+0003$</td>
</tr>
<tr>
<td>$+0016$</td>
<td>$+0001$</td>
</tr>
<tr>
<td>$+0024$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$+0032$</td>
<td>$-0013$</td>
</tr>
<tr>
<td>$-0024$</td>
<td>$-0009$</td>
</tr>
<tr>
<td>$+0007$</td>
<td>$-0029$</td>
</tr>
<tr>
<td>$+0001$</td>
<td>$+0013$</td>
</tr>
<tr>
<td>$-0012$</td>
<td>$+0011$</td>
</tr>
<tr>
<td>$+0005$</td>
<td>$-0010$</td>
</tr>
<tr>
<td>$-0002$</td>
<td>$+0010$</td>
</tr>
<tr>
<td>$-0001$</td>
<td>$-0007$</td>
</tr>
<tr>
<td>$+0003$</td>
<td>$+0112$</td>
</tr>
<tr>
<td>$-0052$</td>
<td>$+0011$</td>
</tr>
<tr>
<td>$-0013$</td>
<td>$+0011$</td>
</tr>
<tr>
<td>$+0010$</td>
<td>$-0010$</td>
</tr>
<tr>
<td>$-0009$</td>
<td>$+0010$</td>
</tr>
<tr>
<td>$-0008$</td>
<td>$+0009$</td>
</tr>
<tr>
<td>$+0007$</td>
<td>$-0008$</td>
</tr>
<tr>
<td>$+0006$</td>
<td>$+0007$</td>
</tr>
<tr>
<td>$-0005$</td>
<td>$-0006$</td>
</tr>
<tr>
<td>$+0004$</td>
<td>$+0006$</td>
</tr>
<tr>
<td>$-0003$</td>
<td>$+0006$</td>
</tr>
<tr>
<td>$+0002$</td>
<td>$-0002$</td>
</tr>
<tr>
<td>$+0001$</td>
<td>$-0001$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$+0000$</td>
</tr>
<tr>
<td>$-0000$</td>
<td>$-0000$</td>
</tr>
</tbody>
</table>
```
+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  +0363  +0422  +0843  +0468  +0429  +0512  +0332  +0355  +0775  +0256
+0463  +0423  +0313  +0593  +0216  +0395  +0399  +0399  +0399  +0399  +0399
+0000
+0c16  +1241  +0509  +1102  +0761  +0725  +0995  +0764  +0681  +0864  +0902  +0783
+0075  +0728  +0634  +0761  +1013  +0594  +1030  +0871  +0437  +1340  +0646  +0771
+1129  +0536  +0977  +0687  +0949  +0520  +0853  +1569  +0155  +1276  +0878  +0532
+1091  +0550  +0648  +1040  +0384  +0900  +0900  +0900  +0900  +0900  +0900
+0000
+0373  +0337  +0569  +0437  +0484  +0316  +0519  +0374  +0431  +0507  +0351  +0485
+0391  +0394  +0466  +0431  +0356  +0442  +0423  +0337  +0448  +0395  +0539  +0328
+0359  +0406  +0510  +0307  +0372  +0406  +0413  +0437  +0326  +0395  +0427  +0326
+036c  +0506  +0279  +0451  +0424  +0362  +0399  +0399  +0399  +0399  +0399  +0399
+0000
+0600  +0713  +0818  +0866  +0927  +0833  +0660  +0949  +0606  +0643  +0703  +0646
+0691  +0687  +0893  +0654  +0650  +0928  +0750  +0797  +0695  +0823  +1077  +0800
+0653  +0696  +0677  +0663  +0758  +0905  +0747  +0802  +0603  +0802  +0952  +0696
+0770  +0613  +0856  +0755  +0855  +0677  +0704  +0784  +0764  +0764  +0764  +0764
```
Use of Prediction XV

It is necessary to prepare a tape containing
the operators and a tape containing the traces $x_i$, $y_i$,
$z_i$, $u_i$. 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 Sil 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
- **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
- **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 \(x_1, y_1, z_1, \) and \(u_1\) is prepared as a separate parameter and then the four parameters are combined into one long one. The form of each is identical.

<table>
<thead>
<tr>
<th>Octal Address</th>
<th>Octal Address</th>
<th>Octal Address</th>
<th>Octal Address</th>
</tr>
</thead>
<tbody>
<tr>
<td>1054</td>
<td>1054</td>
<td>1054</td>
<td>1054</td>
</tr>
<tr>
<td>1055</td>
<td>1055</td>
<td>1055</td>
<td>1055</td>
</tr>
<tr>
<td>1056</td>
<td>1056</td>
<td>1056</td>
<td>1056</td>
</tr>
<tr>
<td>2037</td>
<td>2037</td>
<td>2037</td>
<td>2037</td>
</tr>
</tbody>
</table>

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 \(x_1, y_1, z_1, u_1\).

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.

<table>
<thead>
<tr>
<th>Octal Address</th>
<th>Contents</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1053</td>
<td>+N</td>
<td>( N = \text{no. of operators on tape} )</td>
</tr>
<tr>
<td>1054</td>
<td>+.XXXX</td>
<td>Ident. no. for first operator</td>
</tr>
<tr>
<td>1055</td>
<td>-0,-1,-2 or -3</td>
<td>-0 if ( x_1 ), -1 if ( y_1 ), -2 if ( z_1 ), -3 if ( u_1 )</td>
</tr>
<tr>
<td>1056</td>
<td>+k</td>
<td>( a \times 10^{-3} )</td>
</tr>
<tr>
<td>1057</td>
<td>+M</td>
<td>( a \times 10^{-1} ) Constants for ( x_1 )</td>
</tr>
<tr>
<td>1060</td>
<td>( \pm XXXX )</td>
<td>( a \times 10^{-1} ) for ( z_1 )</td>
</tr>
<tr>
<td>1061</td>
<td>( \pm XXXX )</td>
<td>( b \times 10^{-1} ) Constants for ( y_1 )</td>
</tr>
<tr>
<td>1062</td>
<td>( \pm XXXX )</td>
<td>( c \times 10^{-1} ) Constants for ( u_1 )</td>
</tr>
<tr>
<td>1070</td>
<td>.</td>
<td>( d \times 10^{-1} ) Ident. no. for second operator</td>
</tr>
<tr>
<td>1071</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1100</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1101</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1110</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1111</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1120</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1121</td>
<td>.</td>
<td>etc.</td>
</tr>
<tr>
<td>1165</td>
<td>.</td>
<td>etc.</td>
</tr>
</tbody>
</table>

### 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.
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

\[ R = 1 - \frac{I_{\text{min}}}{{I_{\text{var}}}} \]

which is a measure of the goodness of the least squares fit. A sample of the output appears below.

\begin{verbatim}
Variance sum = 0.04953
Minimum sum  = 0.01840
Operator No. -1010
N = 066
n = 050
k = 002
M = 003
T4 predicted
a 10000 = +0292

a \beta 10 = -0000
a \beta 210 = +0565
a \beta 310 = -0000
a \beta 010 = +0005

d \beta 310 = -0341
d \beta 210 = -0117
d \beta 110 = -0410
d \beta 010 = +0708
\end{verbatim}

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, Sil 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.

<table>
<thead>
<tr>
<th>Octal Address</th>
<th>Contents</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1001</td>
<td>+N</td>
<td>First member op. interval</td>
</tr>
<tr>
<td>1002</td>
<td>+n</td>
<td>Length of op. interval</td>
</tr>
<tr>
<td>1003</td>
<td>+o, or -1</td>
<td>-1 if x, not used</td>
</tr>
<tr>
<td>1004</td>
<td>+o, or -1</td>
<td>-1 if y</td>
</tr>
<tr>
<td>1005</td>
<td>+o, or -1</td>
<td>-1 if z</td>
</tr>
<tr>
<td>1006</td>
<td>+o, or -1</td>
<td>-1 if u</td>
</tr>
<tr>
<td>1007</td>
<td>+.XXX</td>
<td>First operator no.</td>
</tr>
<tr>
<td>1010</td>
<td>-0,-1,-2, or -3</td>
<td>-0 if x, -1 if y</td>
</tr>
<tr>
<td>1111</td>
<td>+k</td>
<td>-2 if z, -3 if u</td>
</tr>
<tr>
<td>1112</td>
<td>+M</td>
<td>= mean of pred. series x 10^-3</td>
</tr>
<tr>
<td>1113</td>
<td>+.XXX</td>
<td>= mean of pred. series x 10^-3</td>
</tr>
</tbody>
</table>

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
<table>
<thead>
<tr>
<th>200</th>
<th>AC 00</th>
<th>000</th>
<th>AC 00</th>
</tr>
</thead>
<tbody>
<tr>
<td>201</td>
<td>AC 00</td>
<td>001</td>
<td>AC 00</td>
</tr>
<tr>
<td>202</td>
<td>AC 00</td>
<td>002</td>
<td>AC 00</td>
</tr>
<tr>
<td>203</td>
<td>AC 00</td>
<td>003</td>
<td>AC 00</td>
</tr>
<tr>
<td>204</td>
<td>AC 00</td>
<td>004</td>
<td>AC 00</td>
</tr>
<tr>
<td>205</td>
<td>AC 00</td>
<td>005</td>
<td>AC 00</td>
</tr>
<tr>
<td>206</td>
<td>AC 00</td>
<td>006</td>
<td>AC 00</td>
</tr>
<tr>
<td>207</td>
<td>AC 00</td>
<td>007</td>
<td>AC 00</td>
</tr>
<tr>
<td>208</td>
<td>AC 00</td>
<td>008</td>
<td>AC 00</td>
</tr>
<tr>
<td>209</td>
<td>AC 00</td>
<td>009</td>
<td>AC 00</td>
</tr>
</tbody>
</table>

**NOTES**

MIT DIGITAL COMPUTER LABORATORY
OCTAL PROGRAM FORM
TITLE: ITERATION INDEX
AUTHOR: SIMPSON
DATE: 5-10-580
TAPE NUMBER: 2416, 2610, 90
<table>
<thead>
<tr>
<th>Column 1</th>
<th>Column 2</th>
<th>Column 3</th>
<th>Column 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data 1</td>
<td>Data 2</td>
<td>Data 3</td>
<td>Data 4</td>
</tr>
<tr>
<td>Data 5</td>
<td>Data 6</td>
<td>Data 7</td>
<td>Data 8</td>
</tr>
<tr>
<td>Data 9</td>
<td>Data 10</td>
<td>Data 11</td>
<td>Data 12</td>
</tr>
<tr>
<td>Data 13</td>
<td>Data 14</td>
<td>Data 15</td>
<td>Data 16</td>
</tr>
<tr>
<td>Data 17</td>
<td>Data 18</td>
<td>Data 19</td>
<td>Data 20</td>
</tr>
</tbody>
</table>

**Notes:**

MIT DIGITAL COMPUTER LABORATORY

OCTAL PROGRAM FORM

TITLE ITERATION INDEX

AUTHOR SIMPSON DATE

TAPE NUMBER 024 - 312
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

\[ \Sigma_{i+N}^{N+n-1} x_i - j y_i \quad j = 0, 1, 2, \ldots, m \]

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 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 1st 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.
Use of 2559 ml

The instructions for use are

Erase storage, Sil 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, Sil switch down
Read in 2559 ml
Read in 2559 pl4
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 pl3
Start at 677

Traveling Correlations

With the aid of tape 2559 pl0 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 \( k \times 10 \) where \( k = 3, 4, \ldots, 44 \). 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 pl0 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)
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.