

THE USE OF ORTHOGONAL POLYNOMIALS IN THE ANALYSIS
OF METEOROLOGICAL DATA

BY

ZILORE LEO SIMOKO MUMBA

Dissertation submitted in support of an application for the
degree of Master of Science, in the Department of Mathematics
University of Zambia.

M E M O R A N D U M

The work of this dissertation was carried out in the Department of Mathematics, in the University of Zambia, and has not been submitted for any other degree or diploma of any examining body. This is the original work of the author, except where otherwise acknowledged in the text.

SUMMARY

This thesis comprises 5 chapters. The first chapter is concerned with the general problem of setting up approximations to arbitrary functions by linear combinations of sets of functions of known structure. The second chapter discusses the mathematical basis of the various approximation criteria, namely L_1 , L_2 (Least squares) and Minimax. Chapter III outlines the method of orthogonal least squares approximation and discusses its advantages over the conventional methods discussed in Chapter II.

The fourth chapter describes a project undertaken by the author as a practical application of orthogonal least squares approximation. The results of fitting the 850 hPa height fields over Southern Africa by bivariate Legendre functions of order 1 to 4, are shown in figures 1 to 10.

Chapter V summarises the work in this dissertation and contains a discussion on the applicability of the method to actual analysis of meteorological data.

ACKNOWLEDGEMENT

I am greatly indebted to professor E.F. Bartholomeusz for his encouragement throughout the process of writing up this dissertation. My thanks are also due to Mrs. Bartholomeusz for going through the list of references to ensure that standard bibliographical format is adhered to. Lastly I would like to thank Mrs. Kajoba for typing the manuscript.

CONTENTS

	PAGE
INTRODUCTION	1
Chapter I	8
1.2 Complete orthogonal sets	10
1.3 The best approximation to a function $f \in H$ in a finite orthonormal set $\Phi = \{\phi_0, \phi_1, \dots, \phi_{n-1}\} \subset H$	11
1.4 Orthogonality of countable (finite or infinite) linearly independent sets of functions	13
1.5a Particular orthonormal sets obtained by the orthogonalization of the linearly independent set x^k	14
1.5b The classical orthogonal polynomials	14
1.6 The basic properties of the family of classical orthogonal polynomials	16
Chapter II	17
2.10 Statement of the problem	17
2.20 Approximation to continuous functions by polynomials. The least squares solution criteria	18
2.30 The case of parameter fit	18
2.31 Estimate of error in Chebyshev's solution	19
2.32 Comments on Taylor's and Lagrange's theorems	21
2.33 The minimax problem	21
2.40 The L_2 norm of least squares approximation	22
2.41 Shortcomings of the method	23

	PAGE
Chapter III	26
3.10 Orthogonal least squares	28
3.11 Advantages of orthogonal least squares	29
3.20 Choice of orthogonal polynomials	31
3.21 Spherical harmonics - global analysis	31
3.22 Orthogonal polynomial functions - local regional analysis	36
3.23 Spherical Orthogonal Functions (SOFs)	40
3.24 Applications of S.O.F.s	46
Chapter IV Illustrative project	48
Chapter V Summary and conclusions	55
Bibliography	58
Figures	

	PAGE
Chapter III	26
3.10 Orthogonal least squares	28
3.11 Advantages of orthogonal least squares	29
3.20 Choice of orthogonal polynomials	31
3.21 Spherical harmonics - global analysis	31
3.22 Orthogonal polynomial functions - local regional analysis	36
3.23 Artificial Orthogonal Functions (A.O.F.s)	40
3.24 Applications of A.O.F.s	46
Chapter IV Illustrative project	48
Chapter V Summary and conclusions	55
Bibliography	58
Figures	

INTRODUCTION

The problem of Meteorology

The particular meteorological problem which forms the basis of this study is one among three closely related meteorological problems. These may be stated, in their simplest terms as follows:-

- (i) The specification of weather data
(such as pressure, temperature, humidity, stream-function and flow characteristics, etc) over all points of a given area, from available data at a set of fixed station points, within or without that area, at a given point in time;
- (ii) the prediction of these weather characteristics from one point in time at a given station, to a future point in time, at the same station;
- (iii) the mixed problem of the extension of weather data in both space (specification) and time (prediction) from space and time data at given stations.

This third problem is the essential problem of weather forecasting in meteorology. The problem being non-linear does not admit theoretical solution in all but a few simple cases and calls for approximate solution procedures in which the error is kept as small ^{as} possible.

Dynamic versus statistical techniques

There are two main approaches to the problem of weather prediction. One is the dynamic approach in which the physical laws governing the behaviour of the atmosphere are investigated, the other is the statistical approach in which the distribution of selected meteorological parameters are investigated, and the present state of the atmosphere is projected into the future on the basis of its statistical behaviour in the past.

In the dynamical method, solutions of the dynamical equations governing the motion of the system from some initial instant ($t=0$) onwards are attempted using certain simplifying assumptions, and lead to conclusions about its subsequent behaviour. In this approach no consideration is given to the behaviour of the system before $t=0$. The dynamical methods have reached a fairly advanced stage of development since Richardson's unsuccessful attempt at the problem in the early 1920's [21]; they now feature among the operational objective weather prediction techniques employed at most major centres in the world.

Both methods are numerical weather prediction methods. They deal with the same problem but appear to have little in common. However, it may be noted that the dynamical method, as practiced, is not entirely free of empirical relations which are called for in the integration of the governing equations. For example,

the geostrophic balance is based on the observed behaviour of the system rather than on pure dynamical theory. Again, any meaningful statistical investigation needs to be based on dynamical considerations. The attempt to predict the future state of the atmosphere rests on the premise that the system is governed by definite physical laws for, if its observed changes were mere chaotic fluctuations, it would be difficult to envisage the success of the method. On the other hand Wiener, as quoted by Lorenz [16], has shown that "if a statistically stationary system is deterministic in the sense that its future state is exactly determined from its present by a governing dynamics, the future of the system may be predicted by linear regression equations, even if the nature of the dynamics is not known". It would then appear that the more predictable the atmosphere may be by dynamical methods, the more predictable it is likely to be by purely empirical statistical methods.

It would be therefore instructive to examine, in the light of the known success of the dynamical method, the reasons in support of the statistical method.

Firstly, the dynamics in its present form can be considered as representing a small portion of the earth-atmosphere system in view of the complex nature of the influences of external factors such

as the earth or sun, on the atmosphere. Secondly, the density of the present observational network is far from satisfactory. Both considerations contribute to the rapid decay with time, of forecast accuracy. Although the purely mathematical difficulties such as truncation errors and inadequacies of the dynamical features of the models are being steadily reduced with the advancement of computational aids, there is little hope for believing that an easy answer to the above two major problems can soon be found. This forms one justification for the promotion of the statistical approach. The other arises from the fact that the dynamical models, presently developed, are found to be less successful in explaining the atmospheric variations in low latitudes. Furthermore, these areas encompass much of the third world where there is an ever diminishing ability to maintain even the present observational networks.

The statistical method has its own drawbacks. Its objective is to find best fit (or best prediction) formulae. When such formulae are based on specific samples it does not necessarily follow that the 'best' formula is also the best for another sample, or indeed for the population. However, since the predictions in

the present context are not concerned with the entire future, it should, in most cases, suffice to make do with suitable formulae and hope to improve these as data availability improves. Secondly, the use of too many predictors (initial data) can be a source of problems. Some of the linear combinations of the predictors may be highly correlated to the predictand in the sample, and as a result, the terms corresponding to this linear combination in the prediction formulae will increase the error when the formula is applied to new data. Therefore, it is of vital importance to reduce to a minimum the number of predictors required.

Thus, whichever technique is employed, an effective means of reducing the volume of initial data is called for. One such means involving the use of orthogonal polynomials, is the central theme of this thesis.

Brief history of the use of polynomial interpolation in the presentation of initial data

In numerical weather prediction, the integration of systems of prediction equations requires a knowledge of the initial values of the field variables at mesh points of a grid. In the early stages of numerical weather prediction, most analysis schemes used around

the world consisted of estimating grid point values of a variable at a given instant, from subjective (i.e. hand drawn) analyses.

Gradually, attempts were made to develop objective analysis techniques towards this end. Several objective analysis methods have been in use during the past 25 years or so. The method of optimum interpolation, which is not discussed in this thesis, has particularly wide application throughout the world. It is the aim of this thesis to investigate whether the representation of meteorological fields by orthogonal polynomials would be of use in providing the initial data field, for numerical prediction.

Among the early attempts in this direction was that of Bushby and Huckle in 1956 [2]. They employed quadratics or cubics of 'best mean square fit' to 500hPa heights over Northwestern Europe and the north Atlantic. Their results turned out to be fairly satisfactory. However, the traditional least squares technique is computationally tedious, especially when polynomials of higher order are taken. Dixon, in 1969 [6], considered the feasibility of representing various meteorological fields in terms of orthogonal polynomial functions, and successfully applied the method to the 300hPa height field over the British Isles,

using bivariate polynomials of degree 1 to 10. His results, employing 8th power approximating polynomials on the 500hPa height field over the same area as that used by Bushby and Huckle, were superior to those obtained by the latter, and proved superior to any subjective analysis, however skilled. In a later paper, Dixon, Spackman, Jones and Frances [7] also considered 6th power three-dimension polynomial fittings of the form $z=f(x,y,p)$ over the same area. Their results were also found to be very satisfactory; the main errors being attributed to the fact that, considering the data volume involved, a 6th power approximation was inadequate for close enough fit. As a rough guide, they proposed that for analysis quality to match the better subjective analysis in two-dimension fitting, the ratio of the number of coefficients in the orthogonal polynomials used, to the number of data values, should be in the range $\frac{1}{4}$ to $\frac{1}{2}$.

The above results point to the desirability of using objective analysis techniques in 2,3 or more dimensions with sets of orthogonal polynomials of increasing degree. This method, quite apart from its computational speed when compared with other methods, has the added advantage that it lends itself well to the incorporation of various statistical tests, during the computation.

CHAPTER I

THE GENERAL ORTHOGONAL POLYNOMIALS-A REVIEW OF THEIR SIGNIFICANCE AND PROPERTIES

This chapter is concerned with the general problem of setting up approximations to any prescribed degree of accuracy, to arbitrary continuous functions in 1, 2 or 3 dimensions, by linear combinations of sets of functions of simpler structure.

That this can be done in the particular case where the data function is known to be continuous and where the approximating set is a set of polynomials, was established by Weierstrass in the following theorem, known as the Weierstrass approximation theorem:

Let f be a continuous real function defined on a closed interval $[\bar{a}, \bar{b}]$, and let ϵ be given.

Then there exists a polynomial p with real coefficients such that $\left| f(x) - p(x) \right| < \epsilon, \forall x \in [\bar{a}, \bar{b}]$. (1.1)

For a proof and discussion of this theorem see Simmons [27].

In any particular instance, the most convenient approximating polynomial sets are those that belong to the class of complete orthonormal sets, for reasons enumerated in the brief discussion below.

The concept of orthonormality of a set of functions in a given function space is based on the admissibility of an inner product operation in that space. The inner product, represented by $\langle f, g \rangle$, of 2 elements f and g of a given function space X , is an operation in that space X that satisfies the following axioms:

Axiom 1: $\langle f, g \rangle \in \mathbb{C}$ (the field of complex numbers)
 \forall all elements f and g of the space.

Axiom 2: $\langle \lambda f, g \rangle = \lambda \langle f, g \rangle, \quad \forall \lambda \in \mathbb{C} \text{ and } f, g \in X.$

Axiom 3: $\langle f + g, h \rangle = \langle f, h \rangle + \langle g, h \rangle, \quad \forall f, g \text{ and } h \in X.$

Axiom 4: $\langle g, f \rangle = \overline{\langle f, g \rangle}$

Axiom 5: $\langle f, f \rangle \geq 0, \quad \forall f \in X; \text{ being } = 0 \text{ if and only if } f = 0 \in X.$

In the context of this dissertation attention is confined to the space $L^2(D)$ of Lebesgue integrable complex or real valued functions on a given finite domain D in \mathbb{R}, \mathbb{R}^2 or \mathbb{R}^3 for which

$\int_D |f|^2 dD < \infty$ in the Lebesgue sense, and for which $\langle f, g \rangle$ is defined as

$$\int_D f(x) \overline{g(x)} dx, \quad x \in D.$$

This can be shown to define an acceptable inner product for this space.

Orthogonal sets in inner product function spaces may be defined as countable or uncountable subsets $F = \{ f \in L^2(D) \mid i \in I \}$ having the property that $\forall i, j \in I, i \neq j, \langle f_i, f_j \rangle = 0$. If in addition $\langle f_i, f_i \rangle = 1, \forall i \in I$, then the set F is said to be orthogonal and normal, or orthonormal.

1.2 Complete orthonormal sets

An orthonormal set X is said to be complete if it is exhaustive or maximal, that is if it is impossible to find a non-zero function in the same set X which is orthogonal to every member of F but does not belong to F . Such complete orthonormal sets, when they are countable, are the most appropriate to the present purpose.

The particular space of interest in the present context is the class of separable $L^2[a,b]$ spaces (-an important class of Hilbert spaces). The special properties of separable Hilbert spaces of complete orthonormal sets used here, are expressed in the following theorems:-

Theorem A: Every non-empty Hilbert space contains a complete orthonormal set.

Theorem B: Such orthonormal sets (for separable Hilbert space) are countable.

Theorem C: Let H be a Hilbert space, and let $\{ f_i \mid i \in N \}$ be an orthonormal set in H . Then the following conditions are

equivalent to one another, in the sense that any one of the listed properties implies the other 3.

- (i) $\{ f_i / i \in N \}$ is complete
- (ii) $x \in H \perp \{ f_i \}, \forall i \in N \Rightarrow x = 0 \in H$
- (iii) if x is an arbitrary vector in H , then

$$x = \sum \langle x, f_i \rangle f_i$$
- (iv) if x is an arbitrary vector in H ,
then $\| x \|^2 = \sum | \langle x, f_i \rangle |^2$.

Theorem D (Bessel's inequality): For any orthonormal subset $L = \{ f_i \} \subset H$ and any $x \in H$.

$$\| x \|^2 = \sum | c_i |^2, \text{ where } c_i = \langle f_i, x \rangle.$$

Proofs of the above theorems are given in Simmons [27]. A proper definition of $L^2[a, b]$ is also given in the above text, page 256, example 3. See also Nicolsky [19], chapter 14.

1.3 The best approximation to a function $f \in H$, in a finite orthonormal subset $\Phi = \{ \phi_0, \phi_1, \dots, \phi_{n-1} \} \subset H$

The mean square error $E(a)$ of any linear combination $f_n(a) = \sum_{i=0}^{n-1} a_i \phi_i$ in Φ used as an approximation to a given function f in H , is defined by

$$E_n(a) = \| f - f_n(a) \|^2 = \langle f - f_n, f - f_n \rangle.$$

The 'best' possible approximation f_n to the given function f in the orthonormal set Φ is achieved by a choice of multiples $(a_0, a_1, \dots, a_{n-1})$ for which $E(a)$ is least.

It can be shown [19] that such a best approximation f_n can always be attained and is in fact, given by

$$f_n = \sum_0^{n-1} c_i \phi_i$$
, where $\langle f, \phi_i \rangle$ are the Fourier components of f in the set Φ . The corresponding least error can be shown to be given by $E = \|f\|^2 - \sum |c_i|^2$, which is ≥ 0 in all cases (c.f. theorem D above).

It must be noted that in the case where Φ is a complete orthonormal set in H , the least error $E=0$ (c.f. theorem D), and the best approximation f_n is, in fact, f itself (c.f. theorem C (iii)).

In the statement given above the convergence of $\sum_0^n c_i \phi_i$ is convergence in H or in the sense of mean square convergence, namely $\|f - f_n\| \rightarrow 0$ as $n \rightarrow \infty$.

In practice, mean square convergence need not necessarily imply pointwise convergence of the sequence f_n to f , namely that $\sum c_i \phi_i = f_n$, $\forall x \in [\bar{a}, \bar{b}]$.

However, there are certain well known conditions of f that would ensure the relation.

$f = \sum c_i \phi_i$, $\forall x \in [\bar{a}, \bar{b}]$ in the pointwise sense. The best known conditions are set out below.

(a) The Dirichlet conditions

$f(x)$ is defined and single-valued, except possibly at a finite number of points in $[\bar{a}, \bar{b}]$.

$f(x)$ is periodic outside $[\bar{a}, \underline{b}]$, with period $2L$.
 $f(x)$ and $f'(x)$ are sectionally continuous in
 (a, b) .

(b) The smoothness conditions

$f(x)$ is piecewise smooth on $[\bar{a}, \underline{b}]$.

$f(x)$ is periodic outside $[\bar{a}, \underline{b}]$, with period $2L$.

1.4 Orthogonalization of countable (finite or infinite) Linearly independent set of functions

Given any countable (finite or infinite) linearly independent set of functions $X = \{x_i / i \in N\}$ in it is possible to systematically replace each function of the set X by a function ϕ_{n-1} a linear combination $(x_0, x_1, \dots, x_{n-1})$ in such a way that the new set $\Phi = \{\phi_i / i \in N\}$ is orthonormal. A special iterative construction procedure of Φ from X (known as the Gram-schmidt procedure) is given by:-

$$\phi_n = x_n - \frac{\langle \phi_{n-1}, x_n \rangle}{\|\phi_{n-1}\|^2} \phi_{n-1} - \dots - \frac{\langle \phi_0, x_n \rangle}{\|\phi_0\|^2} \phi_0$$

with $\phi_0 = x_0$ (choice free)

$$e_n = \phi_n / \|\phi_n\|$$

The set $\{\phi\}$ represents a systematic orthogonalization of the given linearly independent set $X = \{x\}$ and the set $E = \{e_n\}$ represents an orthonormalization of Φ .

1.5a Particular orthonormal sets obtained by the Orthogonalization of the linearly independent set x^n — The classical Orthogonal Polynomials

A useful generalization of the inner product in the case of real valued functions $\in L^2[a, b]$, $\langle f, g \rangle = \int_a^b f(x)g(x)dx$ introduced earlier is the following weighted inner product:-

$$\langle f, g \rangle = \int_a^b \rho(x) f(x) g(x) dx$$

where the given weight function $\rho(x)$ has the following properties:-

- (i) $\rho(x)$ is never - ve on $[a, b]$
- (ii) $\rho(x) > 0$ a.e. (almost everywhere) on $[a, b]$,
e.g. its zeroes, if any, form a countable set on $[a, b]$ i.e. $\int_a^b \rho(x) dx$ exists and is a real number.

In this case:-

$$\langle \phi_i, \phi_j \rangle = \int_a^b \rho(x) \phi_i(x) \phi_j(x) dx$$

By suitable choice of $\rho(x)$ and $[a, b]$, it is possible to develop orthonormal function sets of many kinds. These sets $\{ \phi_i(x) \}$ over various $[a, b]$ and with particular choices $\rho(x)$ are among the most important classes of orthonormal functions.

1.5b The Classical Orthogonal Polynomials

For each prescribed weight ρ and span $[a, b]$ and standardising constant k_n = coefficient of x^n in $p_n(x)$, the set $\{ 1, x, x^2, \dots, x^n \}$ generates a specific system $\{ p_n(x) \}$ of orthogonal polynomials, among which some have been selected for special detailed study by reason of:-

- (i) their frequency of occurrence
- (ii) a definitive property (there exists a choice of 3 alternatives) which they have in common [c.f. $P(1), P(2), P(3)$ p16]

These are the classical orthogonal Polynomials and are listed below:-

Table 1.1

Name	Symbol	Range	Weight	Coefficient c
Legendre	$P_n(x)$	-1 1	1	$2^n (\frac{1}{2})^n / n!$
Jacobi	$P_n^{\alpha, \beta}(x); \alpha > -1, \beta > -1$	-1 1	$(1-x)^\alpha (1+x)^\beta$	$2^{-n} \binom{2n+\alpha+\beta}{n}$
Gegenbauer	$C_n^\lambda(x); \lambda > -\frac{1}{2}$	-1 1	$(1-x^2)^{\lambda-\frac{1}{2}}$	$2^n (\lambda)^n / n!$
Tchebychef(1)	$T_n(x)$	-1 1	$(1-x^2)^{-\frac{1}{2}}$	2^{n-1}
Tchebychef(2)	$U_n(x)$	-1 1	$(1-x^2)^2$	2^n
Laguerre (Generalised)	$L_n^\alpha(x); \alpha > -\frac{1}{2}$	0 ∞	$e^{-x} x^\alpha$	$(-1)^n / n!$
Hermite	$H_n(x)$	$-\infty \infty$	e^{-x^2}	2^n

Notes

- (i) $P_n^{\alpha, \beta}(x)$ and $C_n^\lambda(x)$ include $P_n(x)$, $T_n(x)$ and $U_n(x)$ as special cases.
- (ii) The choice of the standardising constant k_n varies from author to author and is made with some simplification in view.
- (iii) The completeness of the listed sets can be proved.
- (iv) Conditions on the free parameters α , β and λ are stipulated to ensure the integrability of $\rho(x)$ in each case.

1.6 The basic properties of the family of classical orthogonal Polynomials

There are three possible properties common to the orthogonal polynomials any one of which can serve as a defining starting point for the theory of such polynomials. These properties are:-

P(1) Definitive property (Roderique's Identity)

$$P_n(x) = \frac{1}{K_n \rho} D^n(\rho X^n)$$

where K_n is a free scaling factor

ρ is the weight function of the system

X is a polynomial of degree 0, 1 or 2.

and for each n ,

$$\rho X^n = D(\rho X^n) = \dots = D^{n-1}(\rho X^n) = 0$$

at $x=a$ and $x=b$

Under these conditions the system $\{P_n\}$ constitutes a ρ -orthogonal system of polynomials, ($P_n = a$ polynomial of degree n).

P(2) The polynomials P_n of each such system satisfy:

(a) a linear homogeneous differential equation of the form $A(x)W'' + B(x)W' + \lambda_n W = 0$

where $A(x)$ and $B(x)$ are independent of n and λ is independent of X .

(b) a recurrence relation

$$X P_n'(x) = (\alpha + \frac{1}{2} n X'' X) P_n(x) + \beta_n P_{n-1}(x)$$

where α and β can be computed in terms of the set constants.

P(3)

$$h_n = \langle P_n, P_n \rangle$$

$$\frac{(-1)^n K_n n!}{K^n} \int_a^b \rho(x) X^n dx$$

CHAPTER II

THE MATHEMATICAL PROBLEM OF NUMERICAL APPROXIMATION

2.10 Statement of the problem

In practice and specifically in weather forecasting the problem is to represent a continuous function $f(x)$ on a given domain $[\bar{a}, \bar{b}]$, when such a (data) function is presented over the domain in the form of a table of values at given stations, in the form:

$$f(x) = a_0 + a_1x + \dots + a_nx^n + E(a).$$

The error $E(a)$ depends upon the choice of constants a_i . The best approximation is obviously that which makes $E(a)$ a minimum that is global in some acceptable sense, the measure of which is referred to as the goodness of fit.

The situations in which this could arise in meteorological practice are numerous. For example, consider the correlation between the yield of a given agricultural crop Y and some selected weather parameter T . It is possible to fit an expression of the form:

$$Y(T) = a_0 + a_1T + \dots + a_nT^n + E(a). \quad (2.2)$$

This would give the value of Y at any given T (where T may be temperature, etc.). In the two-dimension case it is often useful to fit a polynomial $h(x,y)$ in R^2 to a set of data of a plane geographical domain:

$$h(x,y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + \dots + a_{nk}x^n y^k \quad (2.3)$$

which gives h at any given point (x,y) of the domain. The problem is not restricted to 1 or 2 dimensions, and may be stated in any number of physical dimensions.

2.20 Approximation to continuous functions by polynomials. The 'best' approximation criteria

There are various criteria used for determining goodness of fit of approximating functions. The following are some of the well known criteria;

(i) The L_1 criterion

Find an approximation $f_n(x)$ to $f(x) \in [a,b]$, such that $\int_a^b |f(x) - f_n(x)| dx$ is a minimum (2.4)

(ii) The L_2 (Least squares) criterion

Find an approximation $f_n(x)$ to $f(x) \in [a,b]$ such that $\int_a^b |f(x) - f_n(x)|^2 dx$ is a minimum (2.5)

(iii) The Minimax problem (Tchebychef)

Choose polynomials $f_n(x)$, such that

$$\text{Max}_{x \in [a,b]} |f(x) - f_n(x)| \text{ is minimum} \quad (2.6)$$

2.30 The case of pointwise fit

(i) Taylor approximation near x_0 takes the form

$$f(x) = f(x_0) + f^1(x_0)(x-x_0) + \dots + \frac{f^n(x_0)(x-x_0)^n}{n!} + R_n \quad (2.7)$$

where f^1, f^2, \dots, f^{n+1} exist at $x=x_0$ and the error R_n is given by

$$R_n = \frac{f^{n+1}(\xi)(x-x_0)^{n+1}}{(n+1)!} \quad a < \xi < b \quad (2.8)$$

If f^1, f^2, \dots, f^n are known at $x=x_0$, the Taylor polynomial (2.6) gives $f(x)$ in the neighbourhood of $x_0 \in [a, b]$. If the stations are equally spaced, the derivatives f^1, f^2, \dots, f^n can be replaced by finite-difference formulae.

ii) Lagrange's interpolation problem and solution

The problem

Given the values of f at $n+1$ distinct stations x_0, x_1, \dots, x_n arranged such that $x_0 < x_1 < \dots < x_n$, construct a polynomial $p_n(x)$ of degree n such that

(2.9)

$$f(x_i) = p_n(x_i), \quad \forall i = 0, 1, \dots, n$$

The solution

$$p_n(x) = \sum_{i=0}^n L_i(x) f_i(x), \quad L_i(x) = \prod_{\substack{j=0 \\ j \neq i}}^n \frac{(x-x_j)}{(x_i-x_j)} \quad (2.10)$$

2.31 Estimate of error in Lagrange's solution

Let the error at any point $\alpha \in [a, b]$; $\alpha \neq x_0, x_1, \dots, x_n$ be given by $E(\alpha) = f(\alpha) - p_n(\alpha)$

(2.11)

where $p_n(\alpha)$ is Lagrange's polynomial, and let

$$g(x) = f(x) - p_n(x) - \lambda(x-x_0)(x-x_1) \dots (x-x_n) \quad (2.12)$$

where λ is chosen such that $g(\alpha) = 0$, and let g^1, g^2, \dots, g^n represent successive derivatives of g

$$= \frac{f(\alpha) - p_n(\alpha)}{(\alpha-x_0)(\alpha-x_1) \dots (\alpha-x_n)} = \frac{E(\alpha)}{(\alpha-x_0)(\alpha-x_1) \dots (\alpha-x_n)} \quad (2.13)$$

20/...

g has $n+2$ successive zeroes, $x_0, x_1, \dots, x_i, x_{i+1}, \dots, x_n$

g^1 exists in $[a, b]$ and by Rolle's theorem

g^1 has $n+1$ successive zeroes, $\eta_0, \eta_1, \dots, \eta_n$

g^2 exists ...

g^2 has n successive zeroes, $\zeta_0, \zeta_1, \dots, \zeta_{n-1}$

... ..

g^r has $n-(r-2)$ successive zeroes in $[a, b]$

g^{r+1} exists ...

... ..

g^{n+1} has $n-(n+1-2) + 1$ zero

$$g^{n+1}(x) = f^{n+1}(x) - 0 - \frac{d^{n+1}}{dx^{n+1}} (x^{n+1} \dots)$$

$= f^{n+1}(x) - (n+1)! \text{ has exactly one zero } x = \lambda_\alpha \text{ in } [a, b]$

where λ is chosen as in (2.12).

$$f^{n+1}(x) = (n+1)! \Rightarrow \lambda = \frac{f^{n+1}(x)}{(n+1)!}$$

$$\frac{E(\alpha)}{(\alpha-x_0)(\alpha-x_1)\dots(\alpha-x_n)} = \frac{f^{n+1}(\lambda_\alpha)}{(n+1)!} \quad (2.14)$$

Therefore the error $E(\alpha)$ at any point $\alpha \in [a, b]$ is given by

$$E(\alpha) = f(\alpha) - p_n(\alpha) = \frac{(\alpha-x_0)(\alpha-x_1)\dots(\alpha-x_n) f^{n+1}(\lambda_\alpha)}{(n+1)!} \quad (2.15)$$

2.32 Comments on Taylor's and Lagrange's methods

- i) The error estimates (2.8) and (2.15) above are generally not computable due to the unavailability of both η_α and ζ and the $n+1$ derivatives in the case of tabulated functions. However it is possible to use (2.8) and (2.15) to get an order-of-magnitude of the error.
- ii) The expressions (2.8) and (2.15) give point by point estimates of the error. A global estimate, i.e. a measure of the error over the whole span, would be more useful. In many cases, such an estimate can be found as is shown in the following sections (2.33 and 2.40 below).

2.33 The minimax problem

It may be attempted to keep $E(\alpha)$ small for all $[a, b]$, (see sec. 2.31 above) but this is complicated by the indeterminacy of the last factor in (2.15) where η is essentially dependent on α and f^{n+1} is not explicitly known. Suppose, however, that it is known or it is assumed on physical grounds that f^{n+1} is bounded over $[a, b]$. Then the problem of reducing $E(\alpha)$ over the whole span essentially becomes the problem of choosing the $n+1$ stations $x_i \in [a, b]$ such that $\forall x \in [a, b]$, $(x-x_0)(x-x_1)\dots(x-x_n)$ is minimised, or

$$\max_{x \in [a, b]} \left| (x-x_0)(x-x_1)\dots(x-x_n) \right| \text{ is minimised.} \quad (2.16)$$

This is known as the minimax problem and a solution was found by Tchebychef in the following theorem:

Theorem E

The expression

$$\begin{aligned} \text{Max}_{x \in [a, b]} |(x-x_0)(x-x_1)\dots(x-x_n)| \text{ is least when} \\ \text{the stations } x_0, x_1, \dots, x_n | x \in [a, b] \text{ are chosen} \\ \text{to coincide with the zeroes of the polynomial} \\ T_n(x) = \frac{\cos(n \cos^{-1} x)}{2^n}, \quad -1 \leq x \leq 1 \text{ (Tchebychef's} \\ \text{polynomial of } O(n)). \\ = \frac{\cos n\theta}{2^n}, \quad x = \cos \theta, \quad 0 \leq \theta \leq \pi \quad (2.17) \end{aligned}$$

2.40 The method of least squares approximation

Given a function $f(x)$ (assumed continuous) in tabulated form at a fixed set of data stations (x_0, x_1, \dots, x_n) construct a polynomial

$$f_n(x) = a_0 + a_1 x + \dots + a_n x^n \quad (2.18)$$

for which the mean square error

$$Q(a_0, a_1, \dots, a_r) = \int_a^b |f(x) - f_n(x)|^2 dx \text{ is minimum. (2.19)}$$

A necessary condition at minimum Q is that the set

(a_0, a_1, \dots, a_r) satisfy the conditions

$$\frac{\partial Q(a_0, a_1, \dots, a_r)}{\partial a_k} \Big|_{a=\hat{a}} = 0 \quad ; \quad k=0, 1, \dots, r \quad (2.20)$$

and the mean square error criterion for best fit is then given by the following system of $n+1$ linear equations for the coefficients a_0, a_1, \dots, a_r

$$0 = \frac{\partial Q}{\partial a_k} = 0-2 \times \int_a^b x^k f(x) dx + \sum_{i=0}^r a_i \int_a^b x^{i+k} dx + \sum_{j=0}^r a_j \int_a^b x^{k+j} dx. \quad (2.21)$$

$$0 = 2 \left[\sum_{i=0}^r a_i \int_a^b x^{k+i} dx - \int_a^b x^k f(x) dx \right]$$

or

$$\sum_{i=0}^r a_i \int_a^b x^{k+i} dx = \int_a^b x^k f(x) dx \quad (2.22)$$

The normal system (2.22) can be written in the alternative form

$$h_{ik} a_i = b_k, \quad k, i=0, 1, \dots, r$$

or in matrix notation (2.23)

$$H a = b$$

where $H_{n+1}(a, b) = (h_{ik}) = \int_a^b x^{i+k} dx$ is non-singular by virtue of the existence and uniqueness of the interpolating polynomial, and $b_k = \int_a^b x^k f(x) dx$ can be evaluated by quadrature from its observed values at the data stations. The normal system (2.23) has the unique solution

$$a = H^{-1} b \quad (2.24)$$

where H^{-1} is the inverse of H .

2.41 Shortcomings of the method

The least squares procedure outlined above has the following defects:

- (i) The denser the data the greater the precision in values for a_i and the better the fit, but the more tedious the computation, especially so when polynomials of higher order are taken.
- (ii) There is no objective method of reducing the number of predictors used.
- (iii) The normal equations which arise with the basis $1, x, x^2, \dots, x^n$ and equally spaced data stations x_i involve an approximately Hilbert matrix, which is extremely troublesome,

due to the rapidly diminishing value of the determinant of H as n increases. For the particular case of equally spaced x in $[\bar{a}, \bar{b}] \equiv [0, 1]$, the matrix of coefficients is approximately

$$h_{ik} = \frac{1}{i+k-1} \quad i, k = 0, 1, \dots, n+1$$

$$\begin{pmatrix} 1 & \frac{1}{2} & \dots & \frac{1}{n+1} \\ \frac{1}{2} & \frac{1}{3} & \dots & \frac{1}{2n+1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n+1} & \dots & \dots & \frac{1}{2n+1} \end{pmatrix} \quad (2.25)$$

The ratio of the largest to the smallest term, which is evident in (2.25) increases as n increases, and leads to numerical problems of ill-conditioning in the solution of the normal equations.

For the case of an overdetermined system, i.e. where the number of equations is more than the number of unknowns, the problem (2.23) may be expressed in matrix form by

$$\begin{pmatrix} I_{m \times m} & H_{m \times n} \\ \tilde{H}_{n \times m} & 0_{n \times m} \end{pmatrix} \begin{pmatrix} r_{m \times 1} \\ a_{n \times 1} \end{pmatrix} = \begin{pmatrix} b_{m \times 1} \\ 0_{n \times 1} \end{pmatrix} \quad (2.26)$$

where H is the $m \times n$ matrix of column vectors

f_1, f_2, \dots, f_n

i.e.

$$f_1 = \begin{bmatrix} h_{11} \\ h_{21} \\ \dots \\ h_{m1} \end{bmatrix}, \quad f_2 = \begin{bmatrix} h_{12} \\ h_{22} \\ \dots \\ h_{m2} \end{bmatrix}, \quad \dots, \quad f_n = \begin{bmatrix} h_{1n} \\ h_{2n} \\ \dots \\ h_{mn} \end{bmatrix}$$

\tilde{H} is the transpose of H; I is the $m \times m$ unit matrix;
 0 is an $m \times m$ null matrix; r is an $m \times 1$ residual vector
 which constitutes the error of the approximation, and
 is to be minimised; o is an $n \times 1$ null vector and b is
 an $m \times 1$ vector of the data function.

$$(2.26) \text{ can be put in the form} \quad (2.27)$$

$$(\tilde{H}.H)a = \tilde{H}b$$

and has the solution

$$a = (\tilde{H}.H)^{-1} \tilde{H}b \quad (2.28)$$

The meteorological problem is mostly concerned
 with such overdetermined systems, where we attempt to
 approximate to the field of observations within a small
 area of the chart by, for instance, a quadratic of best
 fit, say

$$h(x,y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 \quad (2.29)$$

where the values of the coefficients a_{ij} which give best
 fit are to be determined. These coefficient values are

given by (2.28) form which the observation vector can then be represented as

$$b = z + r, \text{ where } z=Ha. \quad (2.30)$$

It is sometimes desirable to introduce a weighting matrix W into this scheme. For instance it may be desirable that the fitting be better for some parts of the interval than others. In this case the normal system (2.26) becomes (2.31)

$$(\tilde{H}.W.H)a = \tilde{H}.Wb$$

and has the solution

$$a = (\tilde{H}.W.H)^{-1}\tilde{H}.Wb \quad (2.32)$$

The size of the matrix $H\tilde{H}$ depends entirely on the degree of the expression to be fitted, and grows rapidly as polynomials of higher order are taken. Thus the computation becomes more cumbersome and time consuming, and the problem of ill-conditioning of the matrix $H\tilde{H}$ becomes more serious. As $H\tilde{H}$ grows in size the estimates for a_i become unreliable. Infact it has been found (see [7]) that increasing the degree of the fitting expression does not necessarily bring about an ~~improvement~~ in the fitting expression. This is due partly to the worsening condition of the array.

These considerations lead to the search for approximating polynomials that are represented in a form in which the disadvantages inherent in the above formulations can be minimised. The orthogonal polynomials seem to provide a way out, as is demonstrated in the next chapter.

CHAPTER III

3.10 Orthogonal Least Squares

Consider the set of $n+1$ polynomials $P = \{p_n(x) \mid x \in D; n \in I\}$ where $p_k(x)$ is of degree k in x . Let the fitting function $\phi(x)$ be a linear combination of the $p_i(x)$, say

$$\phi(x) = \sum c_j p_j(x) . \quad (3.1)$$

Now the mean square error defines a function $J(c_0, c_1, \dots, c_r)$
 $= \int_a^b |f(x) - \phi(x)|^2 dx$
of the r variables (c_r) . As in (2.19) above, $J(c_r)$ is quadratic in c_r and the necessary condition for minimum $J(c_r)$ is

$$0 = \frac{\partial J}{\partial c_k} = 0 - 2 \int_a^b p_k(x) f(x) dx + 2 \sum c_j \int_a^b p_j(x) p_k(x) dx \quad (3.2)$$

or the normal system

$$\sum c_j \int_a^b p_j(x) p_k(x) dx = \int_a^b p_k(x) f(x) dx. \quad (3.3)$$

If the set $\{p_n(x)\}$ is orthonormal on the working domain D , the integral on the left side of equation (3.3) reduces to

$$\int_a^b p_j(x) p_k(x) dx = \begin{cases} 1 & j=k \\ 0 & j \neq k \end{cases} \quad (3.4)$$

in this case (3.3) takes the uncoupled form

$$c_j = \int_a^b p_j(x) f(x) dx, \quad j=0, 1, \dots, r \quad (3.5)$$

For this choice of functions, the expression for the mean square error $J(c_r)$ takes the form

$$J(c_r) = \|f\|^2 - \sum c_i^2 \quad (3.6)$$

where the coefficients c_i are the Fourier coefficients of f in the given orthonormal system, defined in sec. (1.3), $J(c)_r$ is necessarily > 0 (Bessel's inequality) and tends to zero as n tends to infinity in the case where the orthonormal set is complete in the function space employed (see e.g. Nicolsky [19], p174).

3.11 Advantages of orthogonal least squares

Orthogonal least squares as described above removes many of the difficulties enumerated in sections (2.32) and 2.41) above, thus:

- (i) The computation is made easier by use of orthogonal sets of functions, for the reason that the system of simultaneous equations for the coefficients in the previous procedure (sec. 2.40) is now replaced by a system of linear equations for each coefficient in turn.
- (ii) An expression for the error of approximation is known (viz equation 3.6 above) and the magnitude of this error can easily be estimated.
- (iii) At any stage of the computation an estimate is available for the sharpness of the approximation namely the residual sum of squares.

$$R_k^2 = (f - c_0 \phi_0 - c_1 \phi_1 - \dots - c_k \phi_k)^2 \quad (3.7)$$

where k represents the last stage in the computation. If it is desired to go one degree higher, the new residual sum of squares is

$$R_{k+1}^2 = (f - c_0 \phi_0 - c_1 \phi_1 - \dots - c_k \phi_k - c_{k+1} \phi_{k+1})^2 \quad (3.8)$$

where the coefficients c, c_1, \dots, c_k are exactly the same in both (3.7) and (3.8). From these two equations

$$R_{k+1}^2 - R_k^2 = c_{k+1}^2 \phi_{k+1}^2 \quad (3.9)$$

Hence the contribution of each term to the residual sum of squares can easily be computed, and those terms with little contribution can be ignored.

- (iv) The correlation index between the fitted surface and the actual observations can be computed.
- (v) There is considerable reduction in the storage requirements, from the full set of observation values, to only a few coefficients. The optimum coefficient to data ratio has been estimated to be in the range 1:2 to 1:4 [7]. This means that instead of storing all the n values of any measured parameter for each of the n stations it is sufficient to store a few coefficients needed in the fitting expression.

3.20 Choice of Orthogonal Polynomials

The choice of a particular class of polynomials depends upon the variable being analysed and the domain in which the analysis is carried out. The three main types of polynomial functions used are:

3.21 Spherical Harmonics - Global analysis

This is a two-dimension extension of harmonic Fourier analysis in a spherical co-ordinate system for wave-type variables which may, on physical grounds, be expected to be periodic in space. The application of spherical harmonic analysis to the spatial distribution of pressure, temperature and flow characteristics, especially at the upper levels, is particularly appropriate because the fields of these variables tend to be dominated by long wave undulations that approximate to the sinusoidal shape of the harmonic space functions. In this way the fields can be described quite accurately by a relatively small number of spherical harmonics, in particular by those which represent the longest wavelengths/9

The height z of a surface of constant pressure at a given time can be expressed as a function of latitude ϕ and longitude λ . z can be represented as a series of the form

$$z(\phi, \lambda) = \sum_{m=0}^{\infty} (a_m(\phi) \cos \frac{m\pi\lambda}{L} + b_m(\phi) \sin \frac{m\pi\lambda}{L}) \quad (3.10)$$

This is a usual Fourier analysis at different latitudes ϕ , where m is the number of waves around the earth. Values of the Fourier coefficients a_m and b_m can be computed from the integrals

$$\begin{aligned} a_m(\phi) &= \frac{1}{L} \int_0^{2L} z(\phi, \lambda) \cos \frac{m\pi\lambda}{L} d\lambda \\ b_m(\phi) &= \frac{1}{L} \int_0^{2L} z(\phi, \lambda) \sin \frac{m\pi\lambda}{L} d\lambda \end{aligned} \quad (3.11)$$

Since the values of z are available at a discrete set of points over a grid point network, a_m and b_m need to be evaluated by approximating the integrals in (3.11) along the latitudes by the corresponding sums of values of the heights. The amplitudes of the wave numbers from 1 to m can be calculated from

$$A_m = \sqrt{a_m^2 + b_m^2} \quad (3.12)$$

Radinovic [20] successfully applied this method to the 1000, 850, 700 and 500 hPa heights over the Mediterranean, taking wave numbers 1 to 5.

Expression (3.10) can similarly be expanded in terms of spherical harmonics as

$$z(\phi, \lambda) = \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} (a_{mn} \cos m\lambda + b_{mn} \sin m\lambda) P_n^m(\cos \phi) \quad (3.13)$$

where $P_n^m(\cos \phi)$ are the Legendre functions of order m (Legendre polynomials if $m=0$).

Putting

$$C_n(\lambda) = \sum_{m=0}^{\infty} a_{nm} \cos m\lambda + b_{nm} \sin m\lambda, \quad (3.13) \text{ can be written}$$

$$z(\phi, \lambda) = \sum_{n=0}^{\infty} C_n(\lambda) P_n^m(\mu), \quad \mu = \cos \phi \quad (3.14)$$

The orthogonality relations for this set of functions are

$$\int_{-1}^1 P_n^m(\mu) P_k^m(\mu) d\mu = \frac{2(n+m)!}{(2n+1)(n+m)!} \delta_{nk} \quad ; \quad n, k=0, 1, \dots, \quad (3.15)$$

Multiplying (3.14) by $P_n^m(\mu)$ and integrating from -1 to 1 gives

$$\int_{-1}^1 z(\phi, \lambda) P_n^m(\mu) d\mu = c_n \int_{-1}^1 P_n^m(\mu) P_n^m(\mu) d\mu \quad (3.16)$$

Using the orthogonality condition (3.15), the right hand side reduces to the single term.

$$\frac{2(n+m)!}{(2n+1)(n+m)!} c_n, \quad (3.17)$$

giving

$$c_n = \frac{(2n+1)(n+m)}{2(n+m)} \int_{-1}^1 z(\phi, \lambda) P_n^m(\mu) d\mu \quad (3.18)$$

Hence the coefficients C_n can be determined by a method similar to that discussed for a_m and b_m in (3.11) above. Also from (3.13), a_{mn} and b_{mn} are simply the Fourier coefficients obtained from the expansion of $C_n(\phi)$ in a Fourier series. It follows that a_{mn} and b_{mn} can be obtained from a_m and b_m respectively, as follows:

$$\begin{aligned} a_{mn} &= \int_{-1}^1 a_m(\mu) P_n^m(\mu) d\mu \\ b_{mn} &= \int_{-1}^1 b_m(\mu) P_n^m(\mu) d\mu \end{aligned} \quad (3.19)$$

where $a_m(\mu)$ and $b_m(\mu)$ are those given in (3.11) and can similarly be determined by, e.g. Simpson's or Trapezoidal quadrature, at the latitudes corresponding to μ . The integrals in (3.19) can then be evaluated by use of the Gaussian quadrature formula

$$\int_{-1}^1 f(\mu) d\mu = \sum A_k^n f(\mu_k^n) \quad (3.20)$$

which is of the highest degree of precision $2n+1$ (Krylov, [147], p. 108), and has, for its n nodes, the roots of the Legendre polynomials of order n

$$P_n(x_k^n) = 0. \quad (3.21)$$

The condition of orthogonality corresponding to (3.15) is

$$\sum A_k^n P_n^m(\mu_k^n) P_n^m(\mu_k^n) = \frac{2(n+m)!}{(2n+1)(n-m)!} \delta_{nn'} \quad (3.22)$$

Using (3.20), (3.19) can be re-written as

$$a_{mn} = \sum A_k^n a_m(\mu_k^n) P_n^m(\mu) \quad (3.23)$$

$$b_{mn} = \sum A_k^n b_m(\mu_k^n) P_n^m(\mu)$$

The values of μ_k^n and A_k^n are given in tables, for various values of n , see [147].

Such an analysis enables the various important characteristics of the flow patterns to be studied. These include;

- (i) Space and time variations of the spectral distribution of wave amplitude at various space scales.

- (ii) Contributions of various terms to the total variance of the fields analysed.
- (iii) Spectral distribution of kinetic energy.
- (iv) Propagation of the waves.

For details of the actual analysis, procedure and the results of analyses carried out in specific areas, reference may be made to Radinovic [20].

Due to the large volume of data that would be involved in a global analysis and the uneven distribution of the available data particularly over the southern hemisphere, the above theoretical results can be modified to take account of one hemisphere only. Eliassen and Machenhauer [8] performed a spherical harmonic analysis of the 1000 and 500 hPa heights over the northern hemisphere utilizing the special condition of orthogonality.

$$\int_0^\pi P_n^m(\mu) P_{n'}^m(\mu) d\mu = \delta_{nn'} , \quad n+n' \text{ even} \quad (3.24)$$

for normalized Legendre functions. Utilizing this condition, they were able to employ the expression (3.13) for one hemisphere. For $n+n'$ even, the height of the isobaric surface considered turns out to be symmetric with respect to the equator. For $n+n'$ odd, the surface turns out to be antisymmetric. For this latter case, a_{mn} and b_{mn} are given by

$$a_{mn} = \int_0^1 a_m(\mu) P_n^m(\mu) d\mu \quad (3.25)$$

$$b_{mn} = \int_0^1 b_m(\mu) P_n^m(\mu) d\mu$$

and the Gaussian quadrature formula is

$$\int_0^1 f(\mu) d\mu = \sum_{k=1}^{N/2} A_k^N f(\mu_k^N) \quad (3.26)$$

3.22 Orthogonal Polynomial Functions - Local regional analysis

When the data are given over a geographical domain smaller than a sphere, the analysis is best achieved if more general functions are used, where the sinusoidal form cannot strictly be assumed and where matching of conditions on the boundaries of the domain does not arise. A wide choice of these polynomials exists from the set of the classical orthogonal polynomials enumerated earlier (sec. 1). The method can be applied to any variable which is continuous in space and to any type of surface. The first step in setting up the analysis is to choose a suitable grid of points, bearing in mind two factors. Firstly, the points should be evenly spaced in the co-ordinate directions x and y , though it is not necessary that the grid spacings Δx and Δy be equal. Secondly, the points should be close enough to reproduce all the major features of the pattern. A clear account of the fitting procedure is given in [6]. This is described briefly below, in a specific context.

The given data comprises a set of m observations of pressure - heights h_1, h_2, \dots, h_m measured at the m stations of a grid point network. The set of m height values (h_i) may be represented by an $m \times 1$ row vector

$$\underline{h} = (h_1, h_2, \dots, h_m). \quad (3.27)$$

These values will be expressed in terms of n base vector functions of known analytical form f_1, f_2, \dots, f_n

$$\underline{h} = a_1 \underline{f}_1 + a_2 \underline{f}_2 + \dots + a_n \underline{f}_n \quad (3.28)$$

If $n=m$ the fit is exact. However, this is not necessary as one of the aims of the exercise is to specify the field in terms of the fewest possible functions. Therefore it is assumed that $n < m$ and \underline{r} is the residual vector. The task then is to determine $\underline{a} = (a_1, a_2, \dots, a_n)$, such that \underline{r} is a minimum, in a specific norm.

If the solution is sought using the method of least squares, the problem can be formulated as a partition matrix equation of the form

$$\begin{bmatrix} \underline{I} & \underline{H} \\ \hat{\underline{H}} & \underline{O} \end{bmatrix} \begin{bmatrix} \underline{r} \\ \underline{o} \end{bmatrix} = \begin{bmatrix} \underline{b} \\ \underline{o} \end{bmatrix} \quad (3.29)$$

$\begin{matrix} m \times m & m \times n \\ n \times m & m \times m \end{matrix}$
 $\begin{matrix} m \times 1 \\ n \times 1 \end{matrix}$
 $\begin{matrix} m \times 1 \\ n \times 1 \end{matrix}$

This was considered in chapter II (2.26) and the solution was given in (2.28).

For clarity, consider the two-dimension case in which the base functions are formed by the polynomial set $1, x, y, x^2, xy, y^2, \dots$ and set up the tableau:

$$\begin{array}{c}
 1 \ x \ y \ x^2 \ . \ . \ . \\
 1 \ \underline{x} \ \underline{y} \ \underline{x}^2 \ . \ . \ . \\
 \begin{array}{c} (x_1, y_1) \\ (x_2, y_2) \\ \vdots \\ (x_n, y_n) \end{array} \begin{array}{|c|} \hline \begin{array}{ccccccc} \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{array} \\ \hline \end{array} \begin{array}{c} \\ \\ \\ \\ \\ \\ \\ \end{array} \\
 \end{array} \quad H_{mn} \quad m > n \quad (3.30)$$

where the dots in H are obtained by evaluating each of $1, x, y, \dots$ at the data positions listed on the left. Then the pressure height vector can be represented in the form

$$\underline{h} = a_1 \underline{1} + a_2 \underline{x} + \dots + \underline{r} \quad (3.31)$$

yielding a grid point expression for the component h_{ij}

$$h_{ij} = a_1 + a_2 x_{ij} + a_3 y_{ij} + \dots \quad (3.32)$$

in the form of a polynomial representing the continuous space distribution of the atmospheric variable considered (here pressure-height). However, the determination of coefficients is hampered by the ill-conditioning problem discussed in chapter I.

Using orthogonal fitting, the tableau (3.30) may be replaced by

$$\begin{array}{c}
 (x_1, y_1) \\
 (x_2, y_2) \\
 \dots \\
 \dots \\
 (x_m, y_m)
 \end{array}
 \begin{array}{c}
 \phi_1 \quad \phi_2 \quad \dots \quad \phi_n \\
 \begin{array}{c}
 \cdot \cdot \cdot \cdot \cdot \cdot \\
 \cdot \cdot \cdot \cdot \cdot \cdot \\
 \cdot \cdot \cdot \cdot \cdot \cdot \\
 \cdot \cdot \cdot \cdot \cdot \cdot \\
 \cdot \cdot \cdot \cdot \cdot \cdot
 \end{array}
 \end{array}
 \begin{array}{c}
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot \\
 \cdot
 \end{array}
 \quad m > n \quad (3.33)$$

where $\phi_1, \phi_2, \dots, \phi_n$ are selected orthogonal polynomials in (x, y) . These ϕ satisfy the orthogonality condition

$$\langle \phi_r, \phi_s \rangle = a_r \delta_{rs} \quad (3.34)$$

The normal system (3.29) now takes the form

$$\langle \phi', \phi \rangle \cdot \underline{b} = \langle \phi' \cdot \underline{h} \rangle, \text{ where } \phi' = \text{transpose } \phi \quad (3.35)$$

The orthogonality condition introduced in (3.34) yields the solution

$$\underline{b} = \langle \phi', \underline{h} \rangle / \phi^2, \quad (3.36)$$

and the observation vector \underline{h} is now given by

$$\underline{h} = b_1 \phi_1 + b_2 \phi_2 + \dots + b_n \phi_n + \underline{r}_n \quad (3.37)$$

and the expression for the fitting polynomial is

$$h_i = b_1 \phi_1 + b_2 \phi_2 + \dots \quad (3.38)$$

In summary, the main steps in the fitting procedure are:

- (i) The generation of a sufficient set of orthogonal base functions from the cartesian set $1, x, y, \dots$, indeed from any other linearly independent set such as the classical orthogonal polynomials.

- (ii) The computation of the coefficients of the fitting function

$$b_r = \frac{\langle h_r, \phi_r \rangle}{\langle \phi_r, \phi_r \rangle} \quad (3.39)$$

- (iii) Construction of the fitting function on the required domain.

Note that step (i) takes more computation time than (ii) and (iii) combined. However, this step does not involve the data, and once the functions are constructed for a particular domain, they can be stored and used for different sets of data.

3.23 Empirical Orthogonal Functions (E.O.F.s)

These are not of predetermined form as are spherical harmonics or classical orthogonal polynomials. Their form rather develops as a unique function of the data to which they apply. The shape of each E.O.F. when plotted on a chart on the area from where the data came bears a close resemblance to the anomaly pattern of the variable itself. In this regard the E.O.F.s have a clear physical interpretation.

In trying to specify an arbitrarily large percentage of the total space and time variability of a variable by each of several types of orthogonal polynomials in general, the number of E.O.F.s to do this will be smaller than that of any other type of functions. As the stability of regression prediction depends on the use of as few predictors as possible, as well as on a large data sample, E.O.F.s are the most efficient in reducing the number of

The functions are orthogonal components of the spacewise variation of a field, while the coefficients of different E.O.F.s are orthogonal in time. In contrast with other orthogonal functions the data points need not be evenly spaced in this case. Any geographical arrangement is suitable for E.O.F. analysis, but as a rule E.O.F. analysis is most efficient when the data comes from a fairly uniform distribution of points.

Mathematical derivation of E.O.F.s

The clearest description of the mathematical derivation is that given by Gilman [107] and is reproduced below.

Consider pressure values at N grid points measured at M different times. The observed data at the grid points form an MxN matrix P_{tr} where t represents time and r represents position. Assume that P_{tr} is expressible as a product of two other matrices Q_{ti} independent of r and Y_{ir} independent of t; $i=1,2,\dots,n$. Thus

$$P_{tr} = Q_{ti} Y_{ir} \quad (3.40)$$

A further requirement is that Q and Y satisfy

$$\begin{aligned} Y_{ir} Y'_{ri} &= I_{ii} \\ Q'_{it} Q_{ti} &= D_{ii} \end{aligned} \quad (3.41)$$

where I_{ii} is the identity unit matrix and D_{ii} is a diagonal matrix with all non-diagonal elements equal to zero. Under these conditions

$$\begin{aligned} P'_{rt} P_{tr} &= (Y'_{ri} Q'_{it}) (Q_{ti} Y_{ir}) \\ &= Y'_{ri} (Q'_{it} Q_{ti}) Y_{ir} \end{aligned}$$

so that

$$Y_{ir} (P'_{rt} P_{tr}) Y'_{ri} = Q'_{it} Q_{ti} = D_{ii} \quad (3.42)$$

and with (3.40) and (3.41)

$$Q_{ti} = P_{tr} Y'_{ri} \quad (3.43)$$

Thus on the basis of the assumed decomposition of P_{tr} into a time factor Y_{ir} and a space factor Q_{ti} , subject to (3.42) and (3.43), Y_{ir} turns out to be simply the matrix which diagonalizes $P'_{rt} P_{tr}$ and Q_{ti} is simply given by (3.43).

Observations

- (i) If P_{tr} are generated as departures from the average of the pressures, $P'_{rt} P_{tr}$ is an nxn matrix of covariances among the time series of pressure at the N grid positions; and if the P_{tr} are normalized at each point by dividing by the local σ , $P'_{rt} P_{tr}$ forms an NxN matrix of correlations among the time series of pressure at the N positions.
- (ii) $Q'_{it} Q_{ti}$ forms an nxn diagonal matrix of variances of the time coefficients q_{ti} (in which all the covariances are zero). The total of the variances of q_{ti} is equal to the total time variance of pressure at the N positions.

(iv) The linear independence of the row vectors in Q_{ti} implies that their separate variances represent separate contributions to the total variance of pressure. This makes it possible to rank the E.O.F.s in order of their importance and discard those E.O.F.s with negligible contribution to the total variance. In this way it is possible to filter out small scale and random local variations.

(v) The analysis can be carried out for any field variable such as temperature, pressure-height, etc. as well as time averaged quantities. The procedure is to determine, from the given data, the E.O.F.s corresponding to the time period under consideration over the given area. The result is a set of patterns (y_i) on a decreasing space scale with increasing i , and those y_i which refer to space scales which are not of interest can be identified and discarded.

The method was first described by Lorenz [16] in 1956 in a meteorological context. With the advent of electronic computers which eased the laborious computations involved, numerous papers on the use of E.O.F.s in the statistical processing of meteorological data have appeared.

The determination of Y and Q can be carried out in the simplest possible way by the method of matrix diagonalization due to Jacobi (for a square symmetric matrix). As an illustrative example consider the simple case where A is the 2×2 symmetric matrix

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \quad (3.44)$$

A will be made diagonal by a simple rotation of axes through an appropriate angle θ . From the transformation equations (for a rotation of axes) an orthogonal matrix T can be formed as follows

$$T = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (3.45)$$

Pre-multiply A by T and post-multiply A by T' , to get

$$\begin{aligned} TAT' &= \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \\ &= \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} a_{11}\cos\theta + a_{12}\sin\theta & a_{12}\cos\theta + a_{11}\sin\theta \\ a_{12}\cos\theta + a_{22}\sin\theta & a_{22}\cos\theta - a_{12}\sin\theta \end{bmatrix} \\ &= \begin{bmatrix} a_{11}\cos^2\theta + 2a_{12}\cos\theta\sin\theta + a_{22}\sin^2\theta & a_{12}\cos^2\theta - a_{11}\cos\theta\sin\theta - a_{12}\sin^2\theta \\ -a_{11}\cos\theta\sin\theta - a_{12}\sin^2\theta + a_{12}\cos^2\theta & -a_{12}\cos\theta\sin\theta + a_{11}\sin^2\theta + a_{22}\cos^2\theta \\ & + a_{22}\sin\theta\cos\theta & + a_{12}\sin\theta\cos\theta \end{bmatrix} \\ &= \begin{bmatrix} a_{11}\cos^2\theta + a_{22}\sin^2\theta + 2a_{12}\sin\theta\cos\theta & (a_{22} - a_{11})\sin\theta\cos\theta + a_{12}(\cos^2\theta - \sin^2\theta) \\ (a_{22} - a_{11})\sin\theta\cos\theta + a_{12}(\cos^2\theta - \sin^2\theta) & a_{11}\sin^2\theta + a_{22}\cos^2\theta + 2a_{12}\sin\theta\cos\theta \end{bmatrix} \end{aligned}$$

which matrix will be diagonal if

$$(a_{22} - a_{11})\sin\theta\cos\theta + a_{12}(\cos^2\theta - \sin^2\theta) = 0$$

or

$$\frac{1}{2}\sin 2\theta (a_{22} - a_{11}) = a_{12}\cos 2\theta$$

i.e.

$$\tan 2\theta = \frac{2a_{12}}{a_{22} - a_{11}} \quad (3.47)$$

Equation (3.47) gives the required angle of rotation. In this simple case A was made diagonal by a single rotation through the angle θ . If A is a matrix of large order, the method consists of eliminating the largest off diagonal element of A by a rotation through θ_1 , followed by the elimination of the next largest off diagonal element of A through another rotation through θ_2 , and so on, until the diagonal elements of A are sufficiently dominating.

Recall equation (3.42), which can be put in the form

$$YAY' = D \quad (3.48)$$

where $A = P'_{rt} P_{tr}$ is the $N \times N$ symmetric matrix.

If a_{jk} is the largest off diagonal element of A , then the transformation matrix T is of the form

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & \dots \\ 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \cos\theta_1 & 0 & \dots & -\sin\theta_1 & \dots & \dots & \dots \\ \dots & \dots & \dots & 0 & 1 & 0 & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \sin\theta_1 & 0 & \dots & \cos\theta_1 & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & \dots & \dots & 0 & 0 & 1 & \dots & \dots \end{bmatrix} \quad (3.49)$$

where the angle of rotation θ_1 is defined by

$$\tan 2\theta_1 = \frac{2a_{jk}}{a_{kk} - a_{jj}} \quad (3.50)$$

the first approximation to D is obtained by forming

$$D_1 = T_1 A T_1' \quad (3.51)$$

A repetition of the above procedure, with now the largest off diagonal element of D_1 used to determine the new angle θ_2 gives

$$\begin{aligned} D_2 &= T_2 D_1 T_2' \\ &= T_2 T_1 A T_1' T_2' \end{aligned}$$

In general

$$\begin{aligned} D_r &= T_r D_{r-1} T_r' \\ &= T_r T_{r-1} \dots T_1 A T_1' T_2' \dots T_r' \end{aligned} \quad (3.52)$$

By this process D_r will tend to the diagonal matrix D. The product $T_r T_{r-1} \dots T_1$ will tend to the diagonalizing matrix Y. The diagonal elements of D are the eigenvalues of Y, ordered so that $d_1 > d_2 > \dots > d_m$, or vice versa. The rows of Y are the associated E.O.F.s

3.24 Applications of E.O.F.s.

The most important property of E.O.F.s. is the filtering of data, enabling large data sample storage to be reduced by up to a quarter of the original volume. Apart from this, EOF analysis has application in the field of long range weather forecasting, particularly the following:

(i) Analog forecasting

In this method the hypothesis is that similar anomaly fields of a variable have similar sequels. Thus by comparing a whole library of past records, it is possible to select a particular field which bears the closest resemblance to the current record. If such an exercise were to be carried out subjectively it would no doubt be time consuming and unreliable, and EOFs provide the most reliable filtered fields for analog selection.

(ii) Multivariate regression

Given a field of one variable over a given period and a second contemporary field of another variable over the same area, sets of empirical orthogonal functions for the two variables can be constructed. From the time coefficients q_{ti} of these functions, linear coefficients can be computed relating the two variables. Specifically it is possible to select a suitable circulation index (such as pressure) as a predictand and relate it to some other meteorological parameter to be predicted.

CHAPTER IV

ILLUSTRATIVE PROJECT - The representation of a specific meteorological field (pressure-height) over Southern Africa, in terms of orthogonal polynomials

This chapter describes a project undertaken by the author, as a practical application of orthogonal polynomial analysis. The analysis was carried out on pressure-height (a function of x and y), in geopotential metres above sea level, of the 850 hPa surface over Southern Africa.

In meteorology, the movement of pressure systems has long been established to correspond to the movement of weather systems. Therefore pressure analysis at the earth's surface has been found to be a useful tool in weather forecasting. But before a pressure analysis can be carried out, the pressure values measured at individual stations have to be reduced to a common reference level. Mean sea level is one such convenient level. The main variables which go into the reduction procedure are;

- (i) The pressure at the station level.
- (ii) The mean temperature in a fictitious air column between the station and the mean sea level.

However, for stations at elevations of 500m or more above mean sea level, this reduction introduces unacceptable errors, and it is found convenient in these circumstances to reduce the station level pressure to the height of some constant pressure level, above the

station.

In meteorological convention, certain constant pressure surfaces have been designated as standard surfaces. The 850 hPa surface, at an average of 1500m above sea level, is one such standard surface. It is common practice to reduce station level pressures of stations **at** 500 m or more above sea level to heights of the 850 hPa surface. The reduction formular used is

$$H_{850} = 67.445(273.15 + T) \log_{10} \frac{P_s}{850} + H_s \quad (4.1)$$

where T is the mean temperature of the air column between the station and the 850 hPa surface, H_s is the station height above sea level and P_s is the station level pressure. In practice T cannot be measured and is approximated by T_s , the station level temperature.

Once the reduction is achieved each computed value can be plotted at the corresponding station position on a chart and a contour analysis can be carried out. But, as can ^{be} judged from the preceeding paragraph, the computed values are contaminated by errors of various kind.

This project was concerned with the actual analysis of the computed values into pressure systems, i.e. the fitting of polynomial surfaces of best fit, to the computed

values over the Southern African region. The classical orthogonal polynomials were examined for this purpose, and Legendre polynomials were chosen for the exercise for the simple reason that this polynomial set has weight function unity. The results are presented below, in figures 1- 10.

Discussion of results

Figure 1 shows the 'actual' situation surmised subjectively from the observed station data. The height values shown on the chart were estimated at every 5° latitude/longitude intersection from a chart, drawn by subjective means, with all its attendant error attributes. The main features of the analysed contour pattern, based on this chart are:

- a high pressure centre in the southeast of the area.
- an elongated low pressure centre running WSW-ENE to the north of the high pressure.

Figure 2 shows the approximation by the first degree Legendre polynomial, i.e. the best fitting plane. This involves terms in 1, x and y only. The pattern depicts the highest values to the south-east, gradually falling to the lowest values in the northwest. This has a root mean square error of 1.06 per cent, where r.m.s. error is computed from the formula

$$e_{rms} = \sqrt{\frac{1}{25} \sum \left(\frac{h-h_c}{h} \right)^2} \quad (4.2)$$

where h_c represents the computed field and summation is over all stations i .

Figure 3 is the second degree approximation. Now the contours are more curved. The rms error was computed to be 0.93 per cent for this field.

Figure 4 is the third degree fitting. Both high and low pressure centres take on the right shape almost exactly in their right positions. However, computed values in the low pressure are less than input values by an average of 1.5 per cent, and those in the high centre are higher than input values by a similar margin. The result is a higher gradient from high to low, as compared to the input pattern. The r.m.s. error is 1.9 per cent.

Figure 5 is the fourth degree fitting. There is a further increase in the fitting error in the high and low centres and a further reduction elsewhere. A further heightening of gradient, results. The r.m.s. error still remains 1.9 per cent.

Figure 6 is a fourth degree fitting by Tchebychef polynomials, for comparison purposes. The resultant pattern was similar to that for Legendre polynomials of the same degree (figure 5 above), but the latter gives overall superior results, the r.m.s. error for this being 6.5 per cent.

Figure 7 is the seventh degree Legendre fitting, the highest admissible degree for the selected grid of $5 \times 7 = 35$ points. There is continued increase in the difference between computed and input values in the extremum centres.

To check whether the fitting error was influenced by truncation error, especially for high degree polynomials, the computation was carried out in double precision arithmetic for Legendre polynomials of order 4. The results obtained were exactly the same as those obtained for single precision arithmetic, and are therefore not shown here.

Some Observations

- (i) For a symmetric domain with equally spaced arguments (x and y), it appears appropriate to use the classical orthogonal polynomials. With this choice of functions the orthogonalization process is redundant, and the fitting procedure only involves the normalization of the functions and determination of the fitting coefficients.

- (ii) The results of the project described above shows a theoretically insignificant error. This error is consistently negative in the vicinity of low centres (i.e. lower computed values than input) and consistently positive in high centres. This raises the possibility of incorporating some refinement into the procedure to cater for extremum points, to reduce the error over the field as a whole.
- (iii) It has to be borne in mind that reference to errors is made on the basis of the assumed accuracy of the input data which, as it has already been pointed out earlier on, are themselves contaminated by errors. In a sense the central theme of this thesis is to provide an objective analysis scheme which is mathematically based and to that extent more consistent in the spatial variability of the values. It is just possible that the error originates more in the incompatibility of the input data due to inconsistency of the individual station values.
- (iv) Actual observing stations are not situated at equally spaced points, and the above procedure would be of limited value in terms of operational objective analysis. It would be preferable to obtain, from the raw data over any distribution of station points



(not necessarily equally spaced), a direct objective analysis over that area, and possibly interpolate for values at equally spaced points if need be.

With this consideration in view (i.e. iv) above), the analysis as described above was applied to actual (station value) 850 hPa heights over Zambia and the neighbouring territories. The results are shown in figures 9 (input data = reported station values) and 9 (fourth degree Legendre fitting). Figure 10 shows the distribution of the error (input height minus computed) at each station over the domain. The r.m.s. of the percentage error (as defined above) works out at 0.97. The error in the vicinity of extremum centres is generally consistent with the earlier discussion (figures 3 to 8). This seems to indicate therefore, that for this procedure and choice of functions used, the distribution of points in the x and y directions is immaterial.

CHAPTER V

SUMMARY AND CONCLUSIONS

The purpose of this thesis was to investigate the use of orthogonal polynomials in the analysis of meteorological data. This was with two particular fields of application in mind. The first is the field of objective analysis for dynamic weather forecasting by numerical means, and the second, statistical weather forecasting. The use of orthogonal functions, as compared to other types of functions, is justified by:

- (i) Economy of specification due to elimination of any linearly dependent terms.
- (ii) Ease of computation, afforded by the orthogonality property.
- (iii) The possibility of using recurrence relations in generating the polynomial functions.
- (iv) Ease of incorporating statistical tests, for goodness of fit, if need be.

Three types of orthogonal polynomial functions were discussed. The first of these, the spherical harmonic functions were found suitable for analysis on a spherical scale, especially in the upper levels of the atmosphere, where the fields of some meteorological elements tend to be dominated by wavelike undulations.

It is possible to describe such fields by a small number of space harmonic functions, particularly those which represent the longest wavelengths. Hence the various characteristics (time and space variability) of the dominant wavelengths can be studied.

The second type of functions considered, the classical orthogonal polynomials are suitable for analysis over a geographical domain much smaller than a sphere. A wide choice of such polynomial sets exists, including the classical orthogonal polynomials, though it is not a necessary condition to start with an orthogonal set. These particular functions appear to lend themselves well for use in the objective analysis of various meteorological fields. The actual implementation of their use in the project undertaken and described in the previous chapter has shown that pressure in particular can be fairly accurately represented by orthogonal polynomials of order 3 or 4.

It is the author's opinion, on the basis of the results discussed in chapter iv, page 48 of this thesis, that this particular method can be very successfully applied in objective analysis, though further research is needed in order to develop some refinement procedures to control the values in the extremum centres.

The third types of functions considered, the empirical orthogonal functions, have a special significance for the study of the relation of circulation to climate. The abundance of literature on the subject is testimony to the concensus of the meteorological community as to their usefulness. Equally significant is the use of EOFs in long range weather forecasting. Here, two types of application are possible. One is prediction by analog (or weather types), and the other by multivariable regression. The reduction in the number of variables needed to specify a field, afforded by EOFs, makes them ~~more~~ convenient than any other type of polynomial functions.

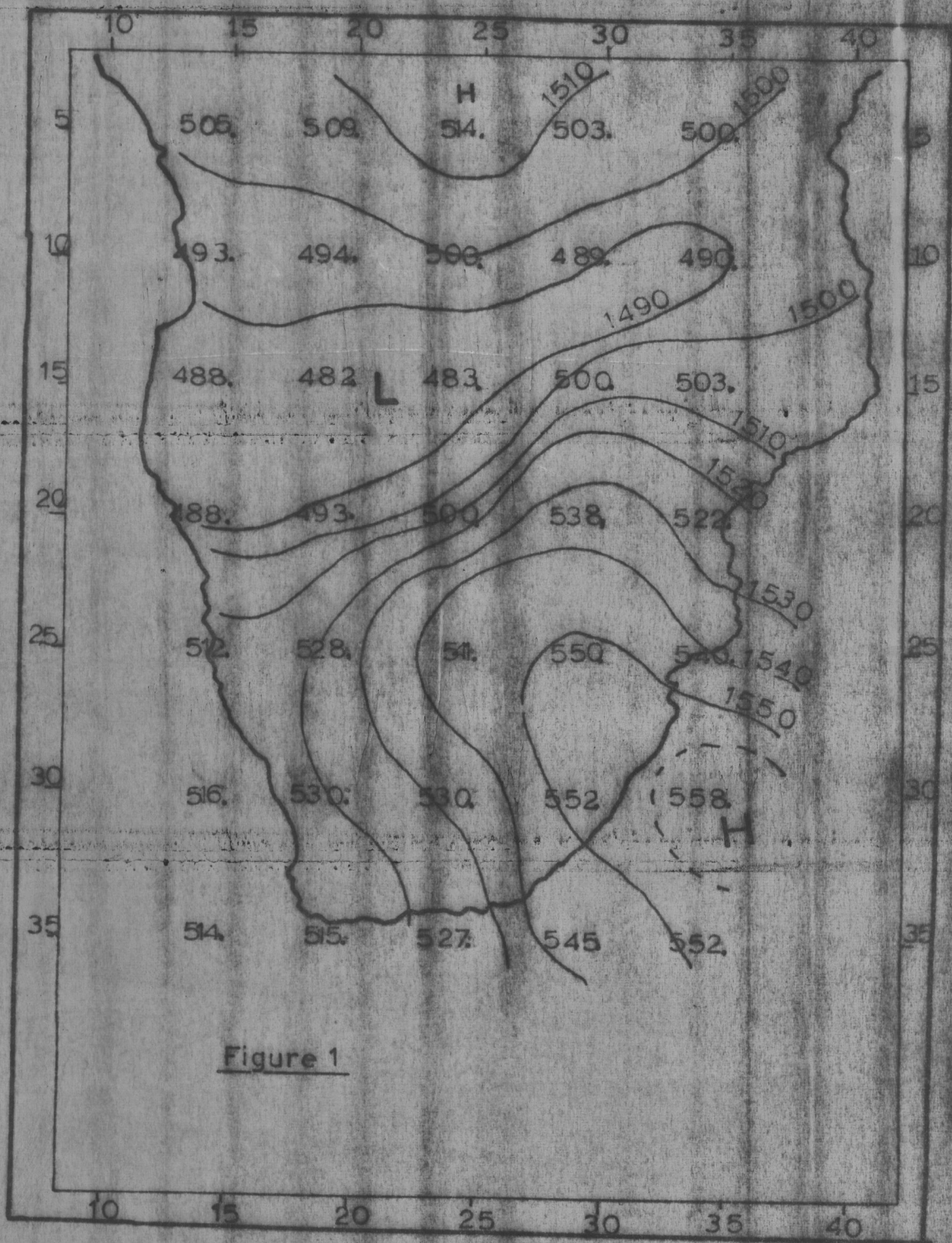
B I B L I O G R A P H Y

1. ALISHOUSE, J.C., CRONE, L.J.,
FLEMING, H.E., VAN CLEEF, F.L. and WARK, D.Q.
A discussion of empirical orthogonal
functions and their applications to
~~vertical temperature profiles.~~ Tellus
XIX, vol. 3(1967), 477-81.
2. BUSHBY, F.H. and HUCKLE, V.M.
Objective analysis in numerical forecasting.
Quart. J. Roy. Met. Soc., (1957), 232-47.
3. CRADDOCK, J.M. and FLOOD, C.R.
Eigenvectors for representing the 500 mb
geopotential surface over the northern
hemisphere. Quart. J. Roy. Met. Soc.,
95(1969), 576-93.
4. CRADDOCK, J.M. and FLINTOFF,
Eigenvector representations of NH fields.
Quart. J. Roy. Met. Soc., 96(1970), 124-9.
5. CRESSMAN, G.P.
An operational objective analysis scheme.
Mon. Wea. Rev., 87(1957) 367-74.
6. DIXON, R. Orthogonal polynomial functions as
a basis for objective analysis. Meteorological
office Sci. paper No. 30, HMSO(1969), 29p.
7. DIXON, R., SPACHMAN, E.A., JONES, I. and ANNE
FRANCIS.
The global analysis of meteorological data
using orthogonal polynomial base functions.
J.Atmos. Sci., Vol. 29, No. 4(1972), 609-18.
8. ELIASSEN, E. and MACHENHAUER, B.
A study of the fluctuations of the atmospheric fl
patterns represented by spherical harmonics.
Tellus XVII, No. 2(1965), 220-38.
9. FORTUS, M.I.
Empirical orthogonal functions in meteorology.
Meteorologiya i Gidrologiya, No. 4(1980), 113-19

10. GILMAN, D.H.
Empirical orthogonal functions applied to 30-day forecasting. Sci. rept. No. 1, Contract No. AF19(604)-1283, Dept. of Meteorology, M.I.T., Cambridge, Mass., 1957, 129p.
11. GRIMMER, H.
The space filtering of monthly temperature anomaly data in terms of pattern, using empirical orthogonal functions. Quart. J. Roy.Met.Soc., 89(1963), 395-408.
12. HANEY, R.L.
Behaviour of the principal harmonics of selected 5-day mean 500mb charts. Mon.Wea. Rev., Vol. 1(1961), 391-6.
13. ISAACSON, EUGENE and KELLER, HERBERT BISHOP.
Analysis of numerical methods. New York: John Wiley and Sons, 1966.
14. KRYLOV, V.I.
Approximate computation of integrals (translated from the Russian). New York: The McMillan Co., 1962.
15. KUTZBACH, J.E.
Large scale features of monthly mean northern hemisphere anomaly maps of sea level pressure. Mon. Wea. Rev., 98, No. 9 (1970), 708-16.
16. LORENZ, E.N.
Empirical orthogonal functions and statistical weather prediction. Sci. rept. No. 2, statistical forecasting project, M.I.T., Dept. of Meteorology Cambridge, Mass, 1957, 151p.
17. LORENZ, E.N.
Prospects for statistical weather forecasting. Final report, statistical forecasting project, M.I.T., Dept. of Meteorology, Cambridge, Mass, 1959, 185p.
18. LORENZ, E.N.
An experiment in nonlinear statistical weather forecasting. Mon. Wea. Rev., 105(1977), 590-602.

19. NICOLSKY, S.M.
A course of mathematical analysis, volume 2.
Translated from the Russian by V.M. Volosov.
Moscow: MIR, 1977.
20. RADINOVIC, D.
Atmospheric waves in the Mediterranean
area on the basis of zonal harmonic analysis.
Tellus, London, 29, No. 3(1977), 213-
21. **RICHARDSON**, L.F.
Weather prediction by numerical process.
London, Cambridge University Press, 1922
22. RINNE, J. and KARHILER, V.
Empirical orthogonal functions of 500mb
height in the northern hemisphere, determined
from a large data sample. Quart. J. Roy. Met. S.
No. 446(1979) 873-84.
23. RINNE, J. and JURVENOVA, S.
Truncation of the EOF series representing
500mb heights. Quart. J. Roy. Met. Soc.,
No. 446(1979), 884.
24. SCHICKEDANZ, P.T. and BOWEN, E.G.
The computation of climatological power
spectra. J. App. Met., Vol. 16(1977),
359-67.
25. SELLERS, W.
A statistical - dynamic approach to
numerical weather prediction of the sea
level pressure field, Sci. rept. No. 3,
statistical forecasting project, M.I.T.,
Dept. of Meteorology, Cambridge, Mass, 1957,
151p.
26. SHOIR, B.
Empirical orthogonal functions applied to
prediction of the sea level pressure field
Sci. rept. No. 3, statistical forecasting
project, M.I.T., Dept. of Meteorology,
Cambridge, Mass, 1957, 88p.
27. SIMMONS, G.F.
An introduction to topology and modern
analysis. International student edition,
New York: McGraw-Hill, 1963.

28. SZERGO, G.
Orthogonal polynomials. American Maths.
Soc. publication, 1939.
29. WEARE, B.C.
Empirical orthogonal analysis of Atlantic
Ocean sea surface temperature. Quart J.
Roy. Met. Soc., 103 (1977), 467-78.
30. W.M.O.
Climatic change. Technical note No. 79,
WMO No. 195, TP100 (1966),
31. W.M.O.
Statistical analysis and prognosis in
Meteorology. Technical note No. 178 WMO
No. TP88 (1966), 16-125.
32. W.M.O.
Symposium on probabilistic and statistical
methods in weather forecasting. Nice,
France, 8-12 September 1980, 235-327.



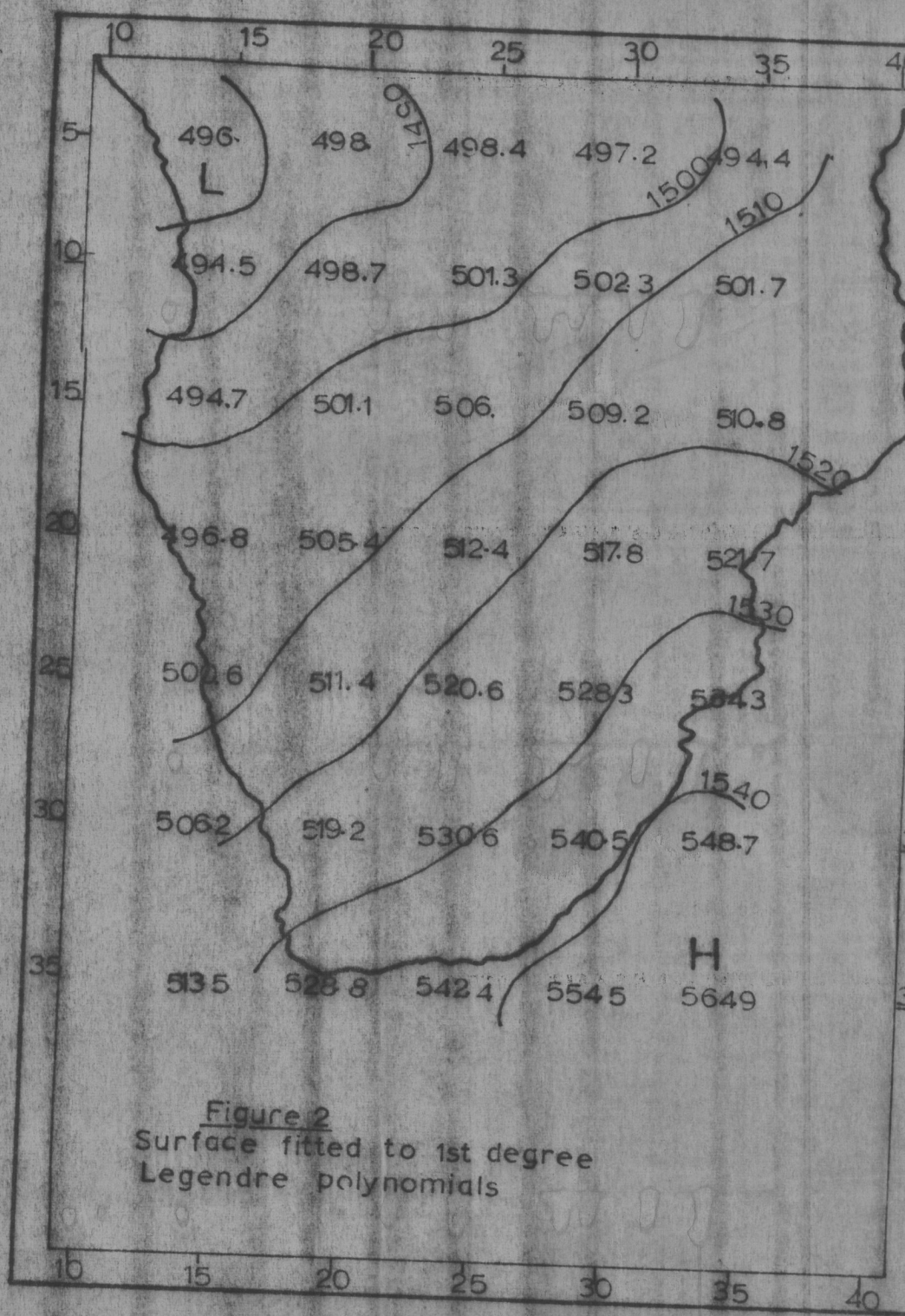
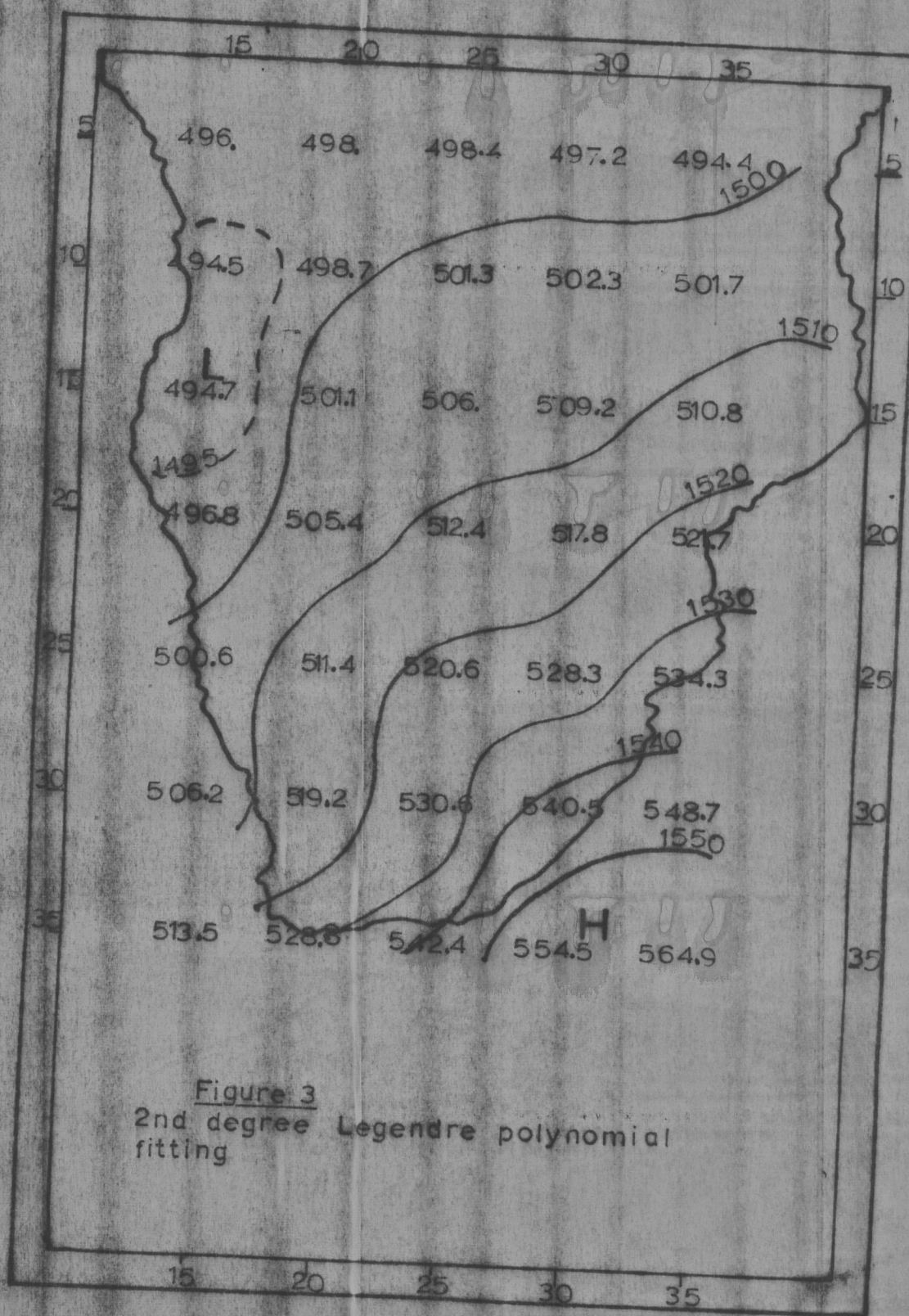
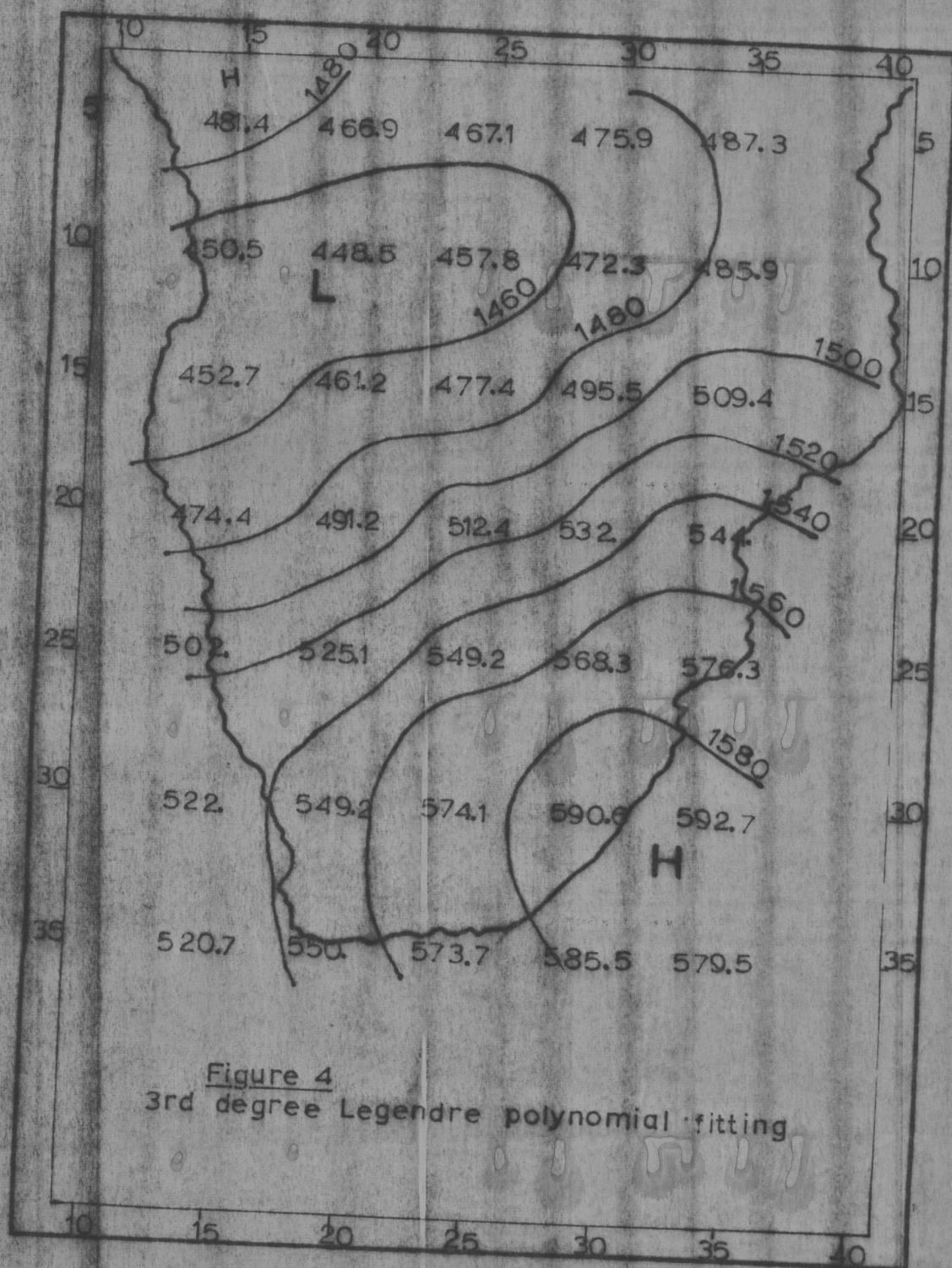


Figure 2
 Surface fitted to 1st degree
 Legendre polynomials





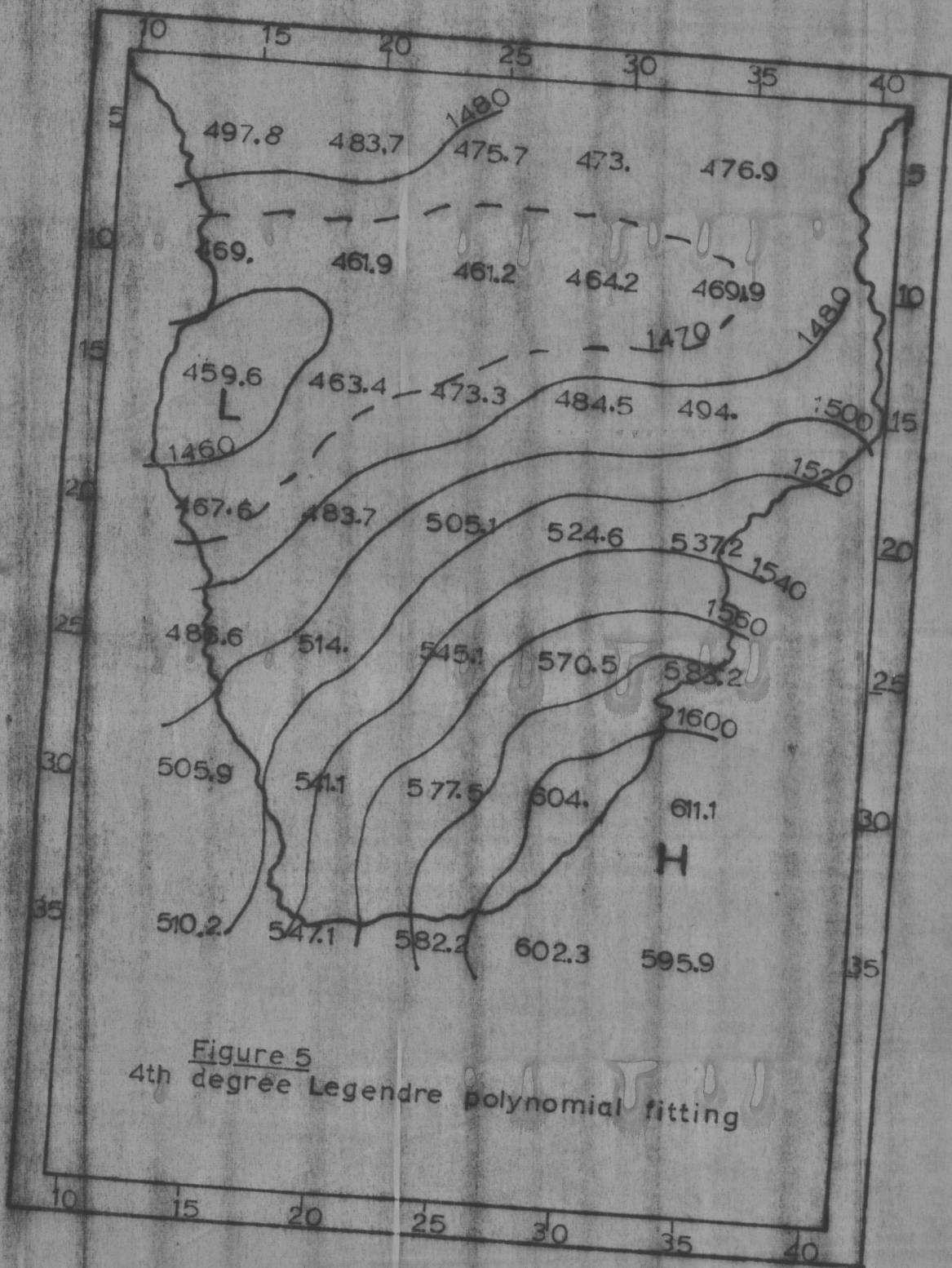


Figure 5
4th degree Legendre polynomial fitting

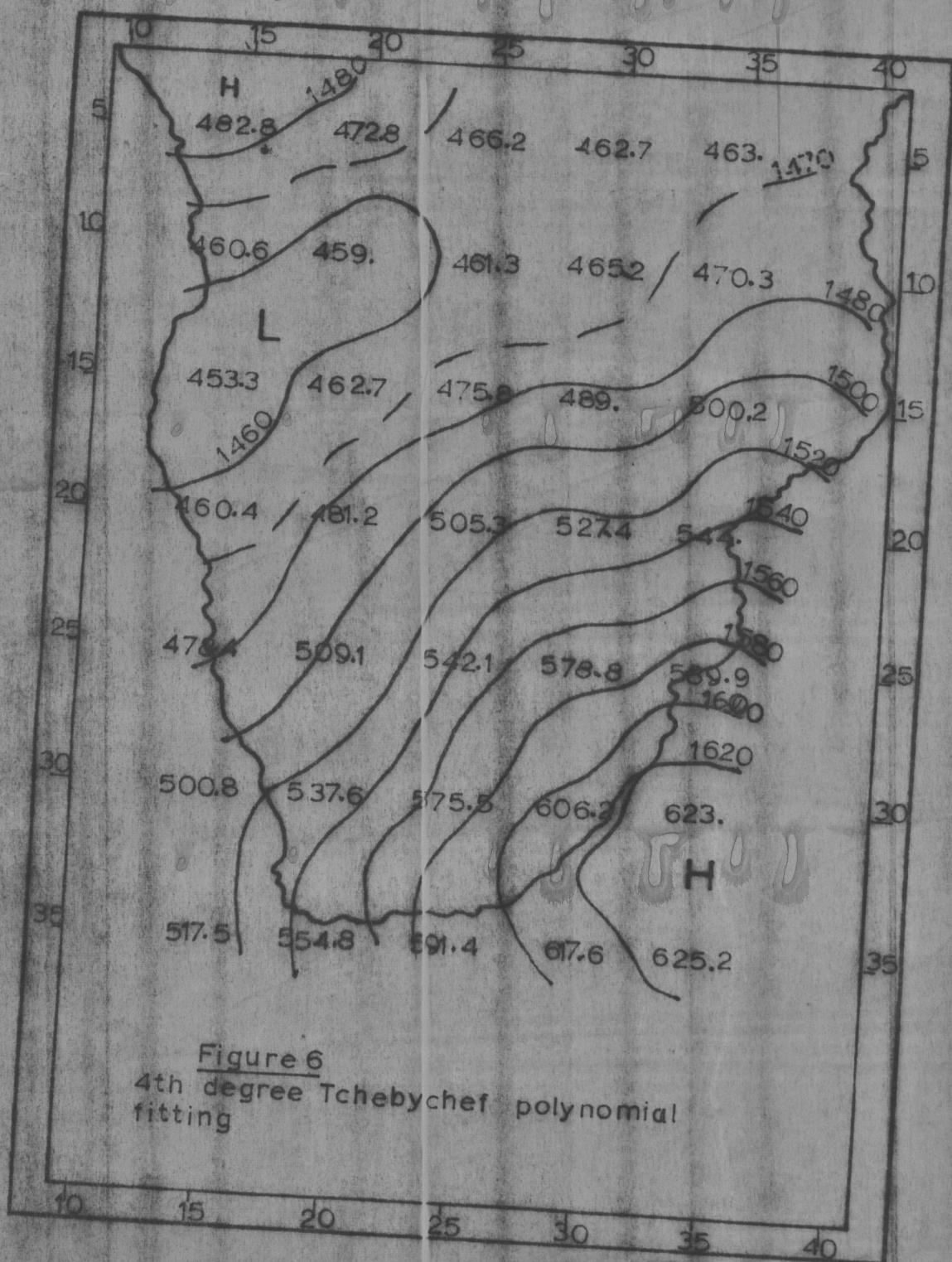
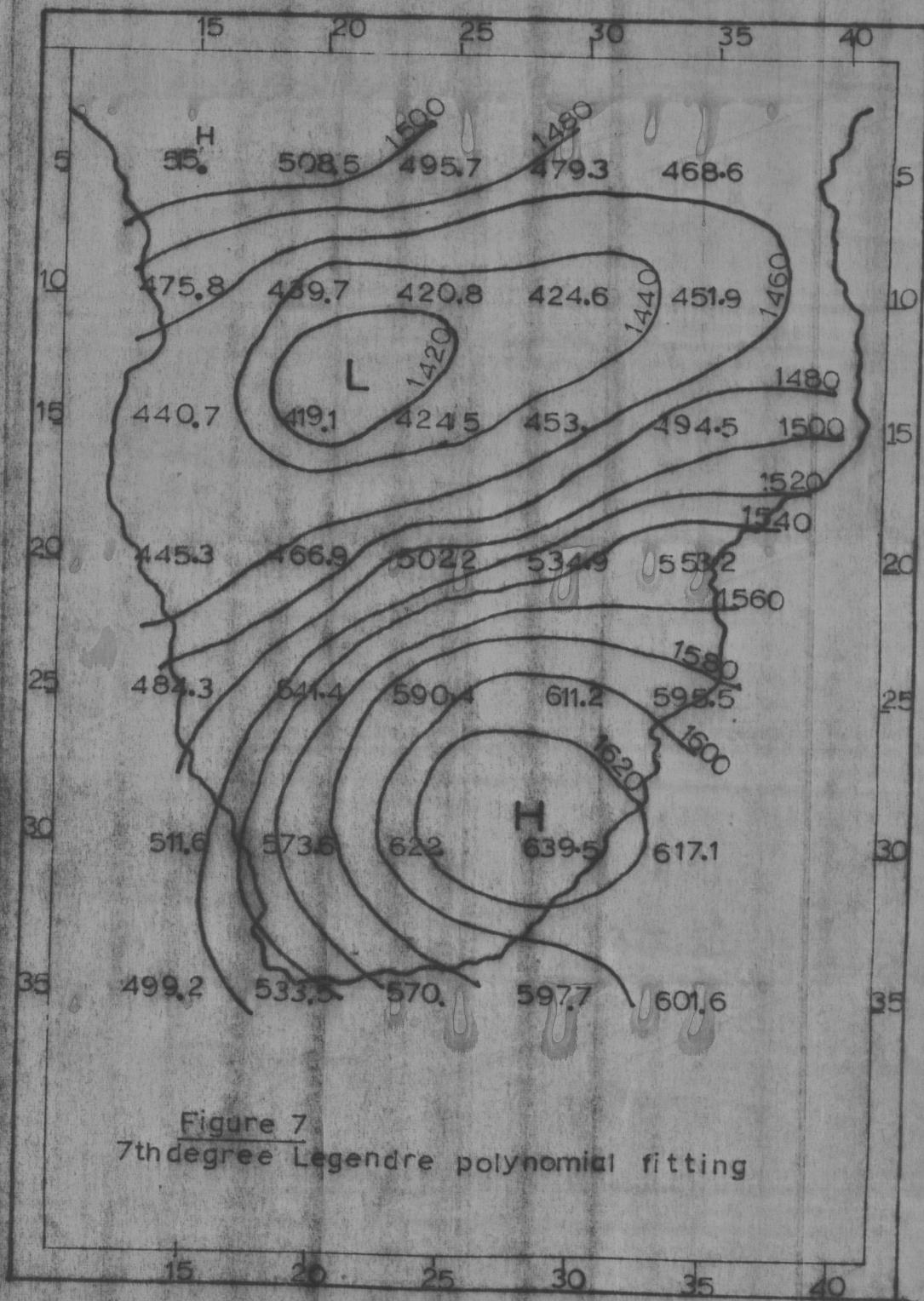


Figure 6
4th degree Tchebychev polynomial
fitting





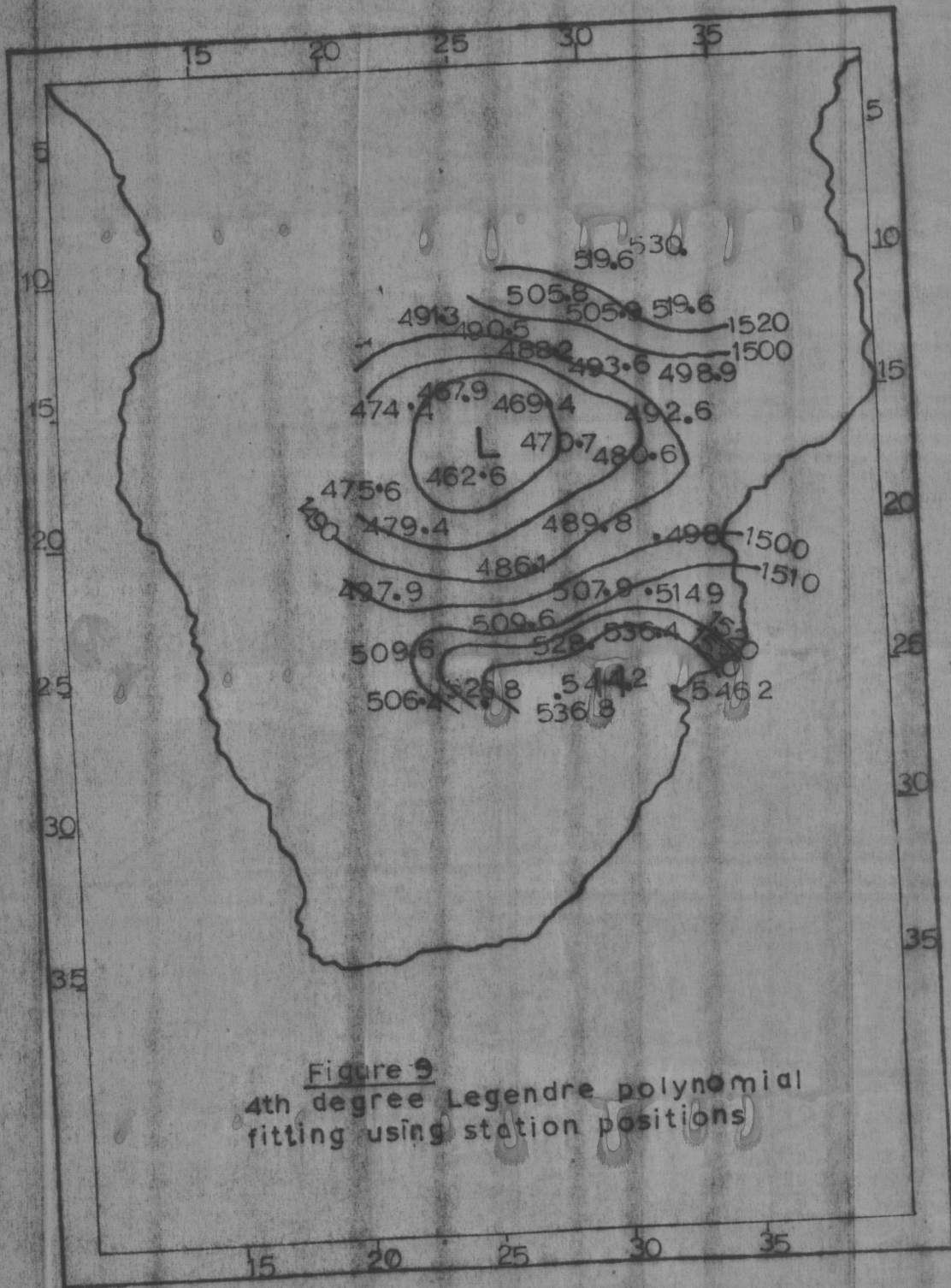




Figure 10
Error distribution for fig. 9