This is to certify that the

dissertation entitled

HOMOTOPY CONTINUATION METHOD FOR
NONLINEAR EQUATIONS

presented by

Mahmoud Mohseni Moghadam

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Mathematics

Tien-Yien Li
Major professor

Date May 4, 1984

HOMOTOPY CONTINUATION METHOD FOR NONLINEAR EQUATIONS

By

Mahmoud Mohseni Moghadam

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Mathematics

1984

ABSTRACT

HOMOTOPY CONTINUATION METHOD
FOR NONLINEAR EQUATIONS

By

Mahmoud Mohseni Moghadam

The essence of the homotopy continuation method is path following. Chapter one of this dissertation is devoted to a detailed discussion concerning certain important aspects of the path following technique. It is shown that the determination of the orientation of the path is a by-product of this computation.

In chapter two we used the homotopy continuation method to determine all roots of a symmetric polynomial system.

Chapter three and four are contributing to the applications of the polynomial systems. In chapter three an algorithm is derived to approximate all real roots of an analytic function in a bounded domain. In chapter four we developed an algorithm to approximate all eigenvalues of a matrix.

Finally in chapter five a special homotopy is constructed to show that there are exactly n distinct

smooth curves connecting n trivial solutions to n eigenpairs of an $n \times n$ given matrix. This homotopy may be used to approximate the eigenpairs of a matrix.

DEDICATED TO

MY PARENTS

AND

MY WIFE

ACKNOWLEDGMENTS

It is a great pleasure to express my sincere appreciation to my thesis advisor Professor Tien-Yien Li. His invaluable guidance, expert advice, stimulating discussion, and useful insights made this work possible.

I would also like to thank Professor Richard Hill for careful reading of this thesis.

Finally, I thank my wife Tahereh Ghasemi for her confidence, patience, and understanding displayed throughout the duration of my graduate study.

TABLE OF CONTENTS

Chapter | Page

CHAPTER ZERO

INTRODUCTION

In the past decade,considerable advances have been made in the problem of obtaining numerical solutions of systems of nonlinear equations. Two closely related methods have become available. (See [24,38] for general references.) The simplicial methods developed by C.E. Lemke and J.T. Howson [30], H. Scarf [41], B.C. Eaves [14,15,16], R. Saigal [40], and others was employed initially for finding the Brouwer fixed point. Their method is based on using a simplicial approximation of the maps as is used in the Sperner's Lemma Proof of the Brouwer fixed point Theorem. R.B. Kellogg, T.Y. Li and J.A. Yorke developed an alternative approach [26,27] for the numerical solution of the Brouwer fixed point Theorem. By a twist of a nonconstructive proof of M. Hirsch [23], they obtained a constructive proof. Given a smooth map F of a ball in $\mathbb{R}^n$ into itself they choose a point p on the boundary of the ball. For almost any choice of p there is a smooth curve Γ_p which leads to a fixed point. The curve can be followed efficiently by computer.

The Brouwer fixed point theorem itself is not used extensively in applications, though similar degree theoretic results are. S.N. Chow, J. Mallet-Paret and J.A. Yorke [4] showed that by using elementary homotopy arguments, numerical methods become available for many problems. (While B.C. Eaves introduced the homotopy idea to the simplicial approach, O.H. Merrill [34] and independently H. Kuhn and J.G. Mackinnon [27] pioneered the sandwich approach.) Their idea is to take any existance proof based on degree theory and to convert it using elementary homotopy into a constructive, computer implementable method for finding solutions. It is shown [4] that the method can be expected to converge with probability one and even is some situations where elementary degree theory is not applicable. This method is more wildly known as "Homotopy continuation method". The essence of the homotopy continuation method is path following. One starts a certain path with a solution which is easily solved, and follows the path until the desired solution is reached. The path is given by the integral of some differential equation under certain regularity assumption, the path is well behaved and can be followed successfully. In fact, simplicial methods and continuation methods are not unrelated as they may initially seem. Both are techniques for following a certain path and are both related to Newton's method.

In chapter one of this thesis, we shall present a detailed discussion concerning certain important aspects of the path following technique. It is shown that the QR decomposition can be applied to the computation of the vector field of the differential equation such that numerical stability can be achieved. It is also shown that the determination of the "orientation" of the path is a by product of this compuation.

The homotopy continuation method is applicable to a considerably large number classes of problems. Among others, its application to locating all the isolated roots of polynomial systems has been most effective and profound. Recently, considerable attention has been given to the problem of finding all solutions of systems of n polynomials in n unknowns by homotopy methods. This problem is important in application and no methods other than homotopy methods are available for finding all the isolated solutions of such systems.

Let

$$\begin{aligned} p_1(z_1,\dots,z_n) &= 0 \\ &\vdots \\ p_n(a_1,\dots,z_n) &= 0 \end{aligned} \tag{1}$$

be n polynomials in n unknowns and write $z = (z_1,\dots,z_n)$, and

$$p(z) = (p_1(z), \ldots, p_n(z)). \tag{2}$$

F.J. Drexler [12,13] and C.B. Garcia & W.I. Zangwill [20,21] showed independently and almost simultaneously that it is possible to find all isolated solutions of $p(z) = 0$. More precisely, it is possible to find a homotopy

$$H : [0,1] \times C^n \to C^n \tag{3}$$

which starts from a trivial set of polynomials

$$\begin{gathered} Q_1(z_1) = 0 \\ \vdots \\ Q_n(z_n) = 0 \,. \end{gathered}$$

That is $Q(z) = 0$, where $Q(z) = (Q_1(z), \ldots, Q_n(z))$,

$$H(0,z) = Q(z), \quad H(1,z) = p(z) \quad ,$$

and has the property that for each isolated zero $\bar{z}$ of p there is a smooth curve Γ of zeros of H in $[0,1] \times C^n$ leading from some zero of Q at $t = 0$ to a zero of $H(1,\bar{z})$.

In [5] S.N. Chow, J. Mallet-paret and J.A. Yorke basically proved Drexler's results but replaced his algebraic geometry arguments with the results involving the generalized Sard's Theorem on which their general homotopy method is based. Garcia and T.Y. Li [19] used those developments to prove a classical theorem of

Noether and Van der Waerden [43] concerning the exact count of the number of solutions of a polynomial system and to give a generalization of the fundamental theorm of algebra. The proof is again free from the algebraic geometry argument used in [37] and is done in such a way that all the isolated solutions can be explicitly calculated. The homotopy used by S.N. Chow, J. Mallet-Paret and J.A. Yorke [5] has been simplified by T.Y. Li [32].

In chapter two of this dissertation, we study a special system of polynomials with strong symmetric properties. The polynomial system is defined by

$$p(z) = (p_1(z), \ldots, p_n(z)) \tag{4}$$

where $z = (z_1, \ldots, z_n)$ and

$$p_i(z) = z_1^i + z_2^i + \cdots + z_n^i, \quad i = 1, 2, \ldots, n. \tag{5}$$

We shall present a special homotopy to exploit the symmetries of this system in such a way that the computations of the curve following can be made much simpler. We also discuss certain techniques to overcome the difficulties when multiple roots are presented in (5).

Chapter three and four are contributing to the applications of the polynomial system (5). In chapter three an algorithm is derived to approximate all the real roots of an analytic function $F(z)$ in a bounded domain. Numerical results up to seven roots are presented to show the effectiveness of this algorithm. In chapter four, we develop an algorithm to approximate all the eigenvalues of an $n \times n$ matrix. In fact, the polynomial system (5) represents the trace formulas for the i<u>th</u> power of a matrix A when z_j's, $j = 1,2,\ldots,n$, are the eigenvalues of A.

In chapter five, we consider the problem of determining all the eigenpairs of a symmetric matrix A. A special homotopy is constructed to show that there are exactly n distinct smooth curves connecting trivial solutions to the desired eigenpairs. Incorporated with sparse matrix techniques, this method may be used to solve eigenvalue problem for large sparse matrices.

CHAPTER ONE

HOMOTOPY CONTINUATION METHOD

§(1.1) Introduction

Let F be a smooth function from $\mathbb{R}^n$ to $\mathbb{R}^n$. In this chapter we shall consider the problem of finding the solution of $F(x) = 0$, by homotopy continuation method. The term "continuation method" is derived from a class of numerical methods dating at least back to E. Lahaye [28,29], and also known as "embedding method". Detailed discussion of these methods can be found in articles by H. Wacker [44], and E. Allgower & K. Georg [2]. One starts with a trivial equation, one to which the solution is obvious and immediately known. Then the system is deformed continuously to $F(x) = 0$. In general, the solution of the trivial system will prescribe, under this deformation, a smooth curve which is connected to the solution of $F(x) = 0$. This curve can be characterized by a solution to an initial value problem of an ordinary differential equation. Our discussion here is limited to following the curve. As a main result, we show that the determination of the orientation of the curve is a byproduct of the computation of the vector field of the ordinary differential equation.

§(1.2) Homotopy

In early 1960's, Davidenko [8], [9] introduced a method of solving $F(x) = 0$ where F is a smooth function from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $H : \mathbb{R}^n \times [0,1] \to \mathbb{R}^n$ be defined as

$$H(x,t) = (1-t)(x-a) + tF(x) \tag{1.2.1}$$

with $a \in \mathbb{R}^n$ given. It is clear that $H(x,0) = x - a$ and $H(x,1) = F(x)$. Suppose

(A): the partial derivative of H with respect to x is always nonsingular

Then by repeat application of the Implicit Function Theorem, there exists a curve $x(t)$, as a function of t, such that

$$H(x(t),t) = 0 \ . \tag{1.2.2}$$

We differentiate (1.2.2) with respect to t, to get the differential equation

$$\begin{aligned} \frac{dx}{dt} &= -H_x^{-1} H_t \\ x(0) &= a \ . \end{aligned} \tag{1.2.3}$$

Thus, finding a zero of $F(x)$ is equivalent to solving the initial value problem (1.2.3), and finding its value at $t = 1$. The assumption (A) is rather strong, and hence the power of Davidenko's method is restricted.

Let us consider the homotopy

$$H : \mathbb{R}^n \times \mathbb{R}^n \times (0,1) \to \mathbb{R}^n \tag{1.2.4}$$

defined by

$$H(x,a,t) = (1-t)(x-a) + tF(x)$$

with $x \in \mathbb{R}^n$, $a \in \mathbb{R}^n$ and $t \in (0,1)$. For a fixed $a \in \mathbb{R}^n$, define

$$H_a : \mathbb{R}^n \times (0,1) \to \mathbb{R}^n \tag{1.2.5}$$

by $H_a(x,t) = H(x,a,t)$.

Given this homotopy system, we define

$$H_a^{-1}(0) = \{(x,t) : H_a(x,t) = 0\} .$$

§(1.3) Existance of a path

We shall use the generalized Sard's Theorem to assure that for almost every $a \in \mathbb{R}^n$, $H_a^{-1}(0)$ contains a smooth curve Γ_a which will lead from a trivial solution to the desired solution.

Definition (1.3.1)

Let U be an open set in $\mathbb{R}^n$, and $f : U \to \mathbb{R}^m$ be a C^1 map, then $y \in \mathbb{R}^m$ is called regular value of f provided

$$\text{range } Df(x) = \mathbb{R}^m \quad \text{for all} \quad x \in f^{-1}(y) \; ,$$

where $Df(x)$ denotes the $n \times m$ matrix of partial derivative of f.

The following theorem may be found in [1].

Theorem (1.3.2) (Generalized Sard's Theorem)

Let $V \subseteq \mathbb{R}^n$, $W \subseteq \mathbb{R}^m$ be open and let

$$G : V \times W \to \mathbb{R}^p$$

be smooth. If $0 \in \mathbb{R}^p$ is a regular value for G, then for almost every $a \in V$ (in the sense of either Baire category or Lebesgue measure), 0 is a regular value for $G_a(\cdot) \equiv G(a,\cdot)$.

For our homotopy defined in (1.2.4) we have the following.

Lemma (1.3.3)

For almost every $a \in \mathbb{R}^n$, zero is a regular value of

$$H_a : \mathbb{R}^n \times (0,1) \to \mathbb{R}^n$$

where H_a is given in (1.2.5).

Proof:

Consider the homotopy $H : \mathbb{R}^n \times \mathbb{R}^n \times (0,1) \to \mathbb{R}^n$, with

$$H(x,a,t) = (1-t)(x-a) + tF(x) \quad .$$

Let $(\bar{x},\bar{a},\bar{t}) \in H^{-1}(0)$, i.e. $H(\bar{x},\bar{a},\bar{t}) = 0$. It is clear that

$$D_a H = -(1-t)I$$

where $D_a H$ is the partial derivative of H with respect to a, and I is the $n \times n$ identity matrix. Since

$$\text{Range } DH(\bar{x},\bar{a},\bar{t}) \supseteq \text{Range } D_a H(\bar{x},\bar{a},\bar{t}) = \mathbb{R}^n \quad ,$$

we conclude that $DH(\bar{x},\bar{a},\bar{t})$ has rank n. Hence 0 is a regular value for H. Thus, by applying the generalized Sard's Theorem we have for almost every $a \in \mathbb{R}^n$, 0 is a regular value of H_a. This completes the proof.

The overall idea is to start from a trivial solution of $H_a(\cdot,0)$ at $t = 0$, and follow the path generated in $H_a(\cdot,t)$ as t goes from zero to one. We hope the trivial solution deforms into the solution of the original system, and hence we would be able to follow the connected path from the trivial system to the solution of $F(x) = 0$.

Of course this is quite an idealized process , and there are a number of difficulties. First of all, in general a path need not exist. Second if one exists, it might be very ill-behaved. In other words the set

$$\{(x,t) : x \in \mathbb{R}^n,\ t \in (0,1),\ H_a(x,t) = 0\}$$

may consist of different solutions, such as isolated points, self-convergings, bifurcations, endless spirals, closed orbit, and smooth paths. But we are interested only in smooth paths (Figure (1.1))

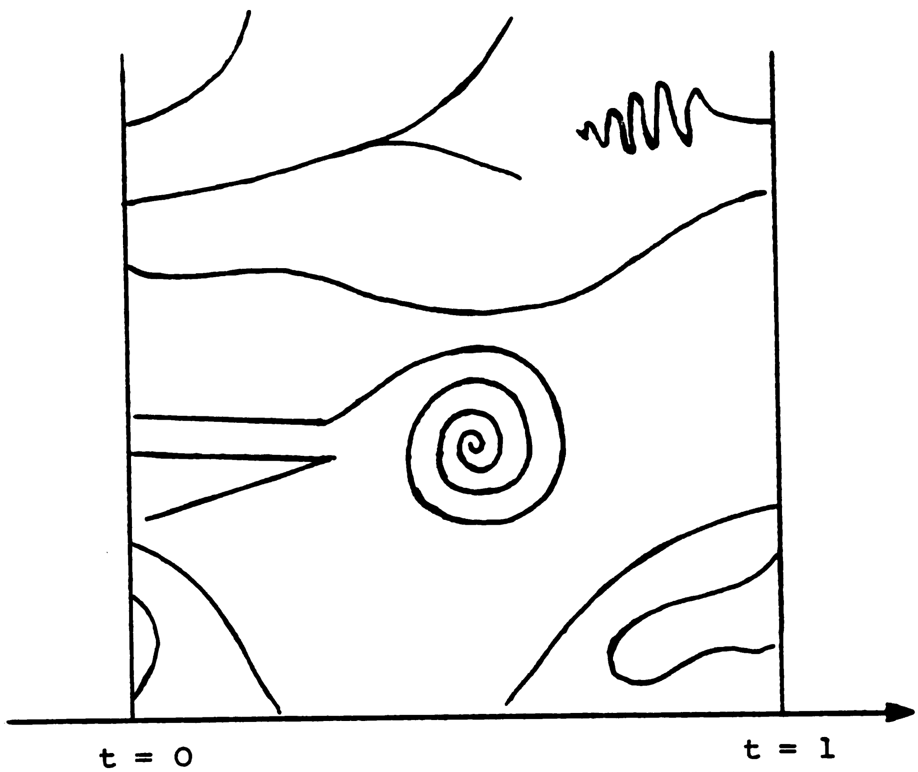


Figure (1.1)

Let a be choosen so that 0 is a regular value for $H_a(x,t)$ (because of Lemma (1.3.3) this can be done with probability one). Then repeated use of the Implicit Function Theorem implies that $H_a^{-1}(0)$ consists of one dimensional manifolds. Detailed discussion for the existance of paths are given by Garcia and Zangwill [22], [20], [21], and Chow, Mallet-Paret and Yorke [4]. Let Γ_a be the component of $H_a^{-1}(0)$ with a as one endpoint. Also let us assume this component is parameterized by s. For notational convenience we refer to $H_a(x,t)$ by $H(x,t)$. Therefore,

$$\begin{aligned} H(x(s),t(s)) &= 0 \\ x(0) &= a \;. \end{aligned} \tag{1.3.6}$$

Differentiation of H with respect to parameter s yields

$$\begin{aligned} H_x(x(s),t(s)) \cdot \dot{x} + H_t(x(s),t(s)) \cdot \dot{t} &= 0 \\ x(0) &= a \;. \end{aligned} \tag{1.3.7}$$

Here H_x and H_t are respectively the partial derivative of H with respect to x and t. The ordinary differential equations (1.3.7) can be written in the following matrix form

$$\begin{aligned} [H_x \quad H_t] \begin{bmatrix} \dot{x} \\ \dot{t} \end{bmatrix} &= 0 \\ \begin{bmatrix} x(0) \\ t(0) \end{bmatrix} &= \begin{bmatrix} a \\ 0 \end{bmatrix} \;. \end{aligned} \tag{1.3.8}$$

The integral solution of this differential equation, namely $(x(s),t(s))$ is a simple curve starting from $(a,0)$. In the next section we carefully examine the movement along this curve.

§(1.4) Movement along the path

We have seen that $H_a^{-1}(0)$ consists of only arcs and closed curves. These curves are the solutions of the ordinary differential equations

$$\begin{gathered} [H_x \quad H_t] \begin{bmatrix} \dot{x} \\ \dot{t} \end{bmatrix} = 0 \\ \begin{bmatrix} x(0) \\ t(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{gathered} \tag{1.4.1}$$

where H_x is an $n \times n$ matrix, H_t is an $n \times 1$ matrix and $\cdot = \frac{d}{ds}$ for some parameter s. For the remainder of this chapter, we will let s be the arc length. Since 0 is a regular value of H, $[H_x \quad H_t]$ is of full rank. Hence kernel of $[H_x \ H_t]$ is one-dimensional, by above the vector $[\dot{x},\dot{t}]^t$ lies in this kernel.

Let $A = [H_x \quad H_t]$ and $\dot{y} = \begin{bmatrix} \dot{x} \\ \dot{t} \end{bmatrix}$. Then (1.4.1) simplifies to

$$\begin{gathered} A\dot{y} = 0 \\ y(0) = \begin{bmatrix} a \\ 0 \end{bmatrix} \end{gathered} \tag{1.4.2}$$

with $\|\dot{y}\|_2 = 1$. The equation $A\dot{y} = 0$ means that $\dot{y}$ is perpendicular to the row space of A. In order to see how $\dot{y}$ can be determined by A, we first give the following definition.

<u>Definition (1.4.1)</u>

A Housholder transformation in $\mathbb{R}^n$ is a matrix of the form

$$U = I - 2vv^t$$

where $v \in \mathbb{R}^n$, $v^t v = 1$, and I is the identity matrix in $\mathbb{R}^n \times \mathbb{R}^n$.

It is easily verified that Housholder transformations satisfy the following properties:

1. $U^t = U$ (symmetric)
2. $U^2 = I$ (involutary)
3. $U^t U = I$ (orthogonal)
4. $\det U = -1$.

Finally the matrix U has the important property that given any two vectors of equal length, x and y, we can find a matrix U such that $Ux = y$. To this end we take

$$v = \frac{x - y}{\|x - y\|_2} . \qquad (1.4.3)$$

By using these properties, a sequence of Housholder transformations, $P_1, P_2, \ldots, P_n$ in $\mathbb{R}^{n+1}$ can be constructed such that

$$P_n \cdots P_1 A^t = R$$

where $R = (r_{ij})$ is an $(n+1) \times n$ upper triangular matrix, that is, $r_{ij} = 0$ for $i > j$. Also P_i's can be choosen in such a way that $r_{ii} > 0$ for $i = 1,2,\ldots,n$. For details see [40,41]. Set

$$Q = (P_n \;\; P_{n-1} \;\; \cdots P_1)^t . \tag{1.4.4}$$

Then Q is an orthogonal matrix and

$$A^t = QR . \tag{1.4.5}$$

Suppose at a certain point $\dot{y}(s)$ and s are known, hence A is known. The following lemma enables us to find $\dot{y}(s + \Delta s)$ for the next point and trace the path by an ordinary differential equation solver.

Lemma (1.4.2)

Let q_{n+1} be the last column of the orthogonal matrix Q, then

$$\dot{y} = \pm q_{n+1} . \tag{1.4.5.1}$$

Proof:

Since $A^t = QR$, and $A\dot{y} = 0$, we have

$$R^t Q^t \dot{y} = 0 .$$

Since matrix R has rank n, we get

$$r_{ii} \neq 0 \quad i = 1,2,\ldots,n . \tag{1.4.6}$$

Suppose $Q^t\dot{y} = (\beta_1, \beta_2, \ldots, \beta_{n+1})^t$, then we have

$$r_{11}\,\beta_1 = 0$$

$$r_{12}\,\beta_1 + r_{22}\,\beta_2 = 0$$

$$\vdots$$

$$r_{1n}\,\beta_1 + r_{2n}\,\beta_2 + \cdots + r_{nn}\,\beta_n = 0 \;.$$

Because of (1.4.6) this system implies

$$\beta_1 = \beta_2 = \cdots = \beta_n = 0 \;.$$

Hence

$$Q^t\dot{y} = (0,0,\ldots,\beta_{n+1})^t \;.$$

So

$$\dot{y} = Q(0,0,\ldots,\beta_{n+1})^t \;.$$

Therefore $\dot{y}$ is a scalar multiple of the last column of Q. Since $\|\dot{y}\|_2 = 1$, we get

$$\dot{y} = \pm\, q_{n+1} \;.$$

In order to determine the orientation of $\dot{y}$, we give the following theorem which can be found in [18].

<u>Theorem (1.4.3)</u>

Let $H : \mathbb{R}^{n+1} \to R^n$ be a C^1 map, and let $z(s) = (z_1(s),\ldots,z_n(s))$ be a C^1 curve in $\mathbb{R}^{n+1}$ satisfying

$$H(z(s)) = 0 \ .$$

Then either

$$\operatorname{sgn} z_i'(s) = \operatorname{sgn} \det H^i(z(s)) \tag{1.4.7}$$

or

$$\operatorname{sgn} z_i'(s) = -\operatorname{sgn} \det H^i(z(s)) \tag{1.4.7.1}$$

for all s, where $z_i'(s) = \frac{dz_i}{ds}$ and H^i is the Jacobian of H with $i\underline{th}$ column deleted.

Applying this theorem to our homotopy, we get either

$$\operatorname{sgn} \dot{t}\,(s) = \operatorname{sgn} \det H_x(x(s),t(s)) \tag{1.4.8}$$

or

$$\operatorname{sgn} \dot{t}\,(s) = -\operatorname{sgn} \det H_x(x(s),t(s)) \tag{1.4.8.1}$$

for all s. However, at $s = 0$, we have $\operatorname{sgn} \det H_x = 1$. We may assume $\dot{t}\,(0) > 0$, therefore (1.4.8) holds for all $s > 0$. If we know $\operatorname{sgn} \dot{t}\,(s)$ the sign in (1.4.5.1) is determined, to determine $\operatorname{sgn} \dot{t}\,(s)$ for any $s > 0$ we prove the following proposition:

<u>Proposition (1.4.4)</u>

Let $Q = (q_{ij})$ be as in (1.4.4), then

$$\operatorname{sgn} \dot{t}\,(s) = (-1)^n \quad \operatorname{sgn}(q_{n+1,n+1}) \ . \tag{1.4.9}$$

<u>Proof</u>:

Let $e_{n+1} \in \mathbb{R}^{n+1}$ be the $(n+1)\underline{th}$ unit vector. That is

$$e_{n+1} = (0,0,\ldots,1)^t$$

then

$$Q^t[A^t\ e_{n+1}] = Q^t[QR\ e_{n+1}]$$
$$= [R\ \ Q^t e_{n+1}]\ .$$

By property of the Housholder transformations

$$\det Q^t = \det P_n \quad \det P_{n-1} \cdots \det P_1 = (-1)^n \quad .$$

Hence

$$\det[R\ \ Q^t e_{n+1}] = \det Q^t \cdot \det[A^t \quad e_{n+1}]$$

$$= (-1)^n \ \det \begin{bmatrix} H_x^t & 0 \\ H_t^t & 1 \end{bmatrix} = (-1)^n \quad \det H_x^t \quad .$$

On the other hand, since R is an upper triangular matrix with $r_{ii} > 0$, we have

$$\operatorname{sgn} \det[R\ \ Q^t e_{n+1}] = \operatorname{sgn}(q_{n+1,n+1}) \cdot \qquad (1.4.10)$$

Therefore

$$\operatorname{sgn} \dot{t} = \operatorname{sgn} \det H_x^t = (-1)^n \quad \operatorname{sgn}(q_{n+1,n+1}) \cdot \quad (1.4.11)$$

From the above discussion we see that in order to follow the curve Γ_a, computation of the vector field

$$\begin{bmatrix} \dot{x}(s) \\ \dot{t}(s) \end{bmatrix}$$

at s can be summerized as follows:

We first compute $H_x(x(s),t(s))$ and $H_t(x(s),t(s))$. Then write

$$A^t = \begin{bmatrix} H_x^t \\ H_t^t \end{bmatrix}$$

as a product of an $(n+1) \times (n+1)$ orthogonal matrix $Q = (q_{ij})$ and an $(n+1) \times n$ upper triangular matrix R. Then $\begin{bmatrix} \dot{x}(s) \\ \dot{t}(s) \end{bmatrix}$ is given by the last column of Q with a possible sign change, and the sign of this vector is given by (1.4.9).

CHAPTER TWO

FINDING THE ZEROS OF AN ANALYTIC FUNCTION

§(2.1) Introduction and notations

Let f be an entire function and R be a bounded domain in the complex plane with a closed boundary C. Let us assume that C does not pass through a zero of $f(z)$. It is well known from the theory of complex variables that

$$\frac{1}{2\pi i} \int_C z^k \frac{f'(z)}{f(z)} = \sum_{i=1}^{n} z_i^k \tag{2.1.1}$$

where $z_i (i = 1,2,\ldots,n)$ are all the zeros of $f(z)$ which lie in R, (a multiple zero is counted according to its multiplicity in its formula).

By using the algorithm explained in chapter three of this dissertation or any other proper algorithm we may find an approximation for

$$S_k = \sum_{i=1}^{n} z_i^k, \qquad k = 0,1,2,\ldots . \tag{2.1.2}$$

Therefore in order to determine the zeros of f we may solve the following system of polynomials.

$$\begin{cases} z_1 + z_2 + \cdots + z_n = s_1 \\ z_1^2 + z_2^2 + \cdots + z_n^2 = s_2 \\ \vdots \\ z_1^n + z_2^n + \cdots + z_n^n = s_n \end{cases} \tag{2.1.3}$$

Delves and Lyness [11,33] developed an algorithm for finding all zeros of $f(z)$ inside R as follows:

i. Let $f(z)$ be an analytic function on and within a given contour C.

ii. Let $z_1, z_2, \ldots, z_k$ be a list of the known zeros of $f(z)$ which have so far been obtained.

iii. Let m, k, and ε be given constants.

iv. Search routine.

The main routine attempts to calculate the number of zeros of $f(z)$ within C using trapezoidal role approximation to contour integral

$$S_0 = \frac{1}{2\pi i} \int_C \frac{f'(z)}{f(z)}\, dz \ .$$

Since the exact result of S_0 is known to be an integer, the accuracy required is low. We only need to determine unambiguously which integer is involved. There are three possible outcomes. These are:

a. It finds that $f(z)$ becomes unduly small on the contour and takes this to imply there are a zero of $f(z)$ close to the contour. In this case the integration method, if continued, would converge slowly. The routine does not continue the integration, but returns control to the search routine, which in turn chooses a different contour.

b. It finds a value S_0. On checking the list of known zeros it finds that q of these lie within the region R, and hence there are $S_0 - q$ unknown zeros within R. If $S_0 - q > M$ it returns control to the search routine.

c. It proceeds as in (b) but finds $S_0 - q \leq M$. It then evaluates the $S_0 - q$ unknown zeros as follows. It evaluates approximations to the sums of power of zeros

$$S_N = \sum_{i=1}^{S_0} z_i^N , \qquad (N = 0,1,\ldots,S_0 - q) \qquad (2.1.4)$$

using trapezoidal approximation to the integral

$$S_N = \frac{1}{2\pi i} \int_C z^N \frac{f'(z)}{f(z)}\, dz .$$

Since the locations of the known zeros $z_1, z_2, \ldots, z_n$ are available the sums $\bar{S}_N$ of the power of unknown zeros are

$$\bar{S}_N = \sum_{i=q+1}^{S_0} z_i^N = S_N - \sum_{i=1}^{q} z_i^N , \quad N = 0,1,\ldots,S_0 - q .$$

Having these numbers, a polynomial of degree $S_0 - q$ may be constructed as follows:

Define

$$\sigma_1 = -(z_1 + z_2 + \cdots + z_{S_0-q})$$

$$\sigma_2 = (z_1 z_2 + \cdots + z_{S_0-q-1}\, z_{S_0-q})$$

$$\vdots \tag{2.1.5}$$

$$\sigma_{S_0-q} = (-1)^{S_0-q}\, z_1 z_2 \cdots z_{S_0-q} \quad .$$

Then by the Newton's identities

$$S_1 + \sigma_1 = 0$$

$$S_2 + S_1\sigma_1 + 2\sigma_2 = 0$$

$$\vdots \tag{2.1.6}$$

$$S_{S_0-q} + S_{S_0-q-1}\,\sigma_1 + \cdots + (S_0 - q)\sigma_{S_0-q} = 0 \quad .$$

One may construct the polynomial

$$P(z) = z^{S_0-q} + \sigma_1\, z^{S_0-q-1} + \cdots + \sigma_{S_0-q} \tag{2.1.7}$$

which has z_i $(i = 1,2,\ldots,S_0 - q)$ as roots. This polynomial is solved by using the polynomial root finding subroutine.

In the above algorithm M is generally choosen to be 5. Although this is a big restriction, it is necessary due to the highly sensitive dependence of high degree polynomial on its coefficients (cf: Wilkinson's example [45]). In other words M cannot be choosen too large in order to avoid the ill-conditioning of the polynomial (2.1.7).

Professor Li [32] recently solved this problem by homotopy continuation method directly without forming the polynomial (2.1.7). Hence he removed the search routine as well as the step (b) from the procedure.

Here we shall present a special homotopy to exploit the symmetries of this system in such a way that the computation of the curve can be made much simpler. That is we will utilize the strong symmetry structure of (2.1.3) to write the corresponding differential equations explicitly in such a way that it will be ready to be solved by any ordinary differential equations solver. We also generalize the problem in the sense that the assumption "f(z) must have single roots" is relaxed. Namely, in section (2.5) we discuss certain techniques to overcome the difficulties when multiple roots are presented in (2.1.3).

§(2.2) Method and Homotopy

Let $Q : C^n \to C^n$ be a polynomial system, and $Q = (Q_1, \ldots, Q_n)$. By a polynomial system, we mean each term of Q_i is of the form

$$a z_1^{r_1} z_2^{r_2} \cdots z_n^{r_n} \tag{2.2.1}$$

where a is a complex number, z_i a complex variable, and

r_i a nonnegative integer. For each term (of the form (2.2.1)) in Q_i, consider the sum $r_1 + r_2 + \cdots + r_n$. Let q_i be the maximum such sum in Q_i. We assume $q_i > 0$, for all i. We call q_i the degree of Q_i. With these notations we state the following theorem.

Theorem (2.2.1)

Let $Q : C^n \to C^n$ be a polynomial system with degree $Q_i = q_i$ $(i = 1,2,\ldots,n)$. Let $Q^0(z) = (Q_1^0(z), Q_2^0(z), \ldots, Q_n^0(z))$ where $Q_i^0(z)$ consists of the terms in $Q_i(z)$ with degree q_i. If $Q^0(z)$ has only the trivial solution $z = 0$, then $Q(z) = 0$ has

$$\prod_{i=1}^{n} q_i \qquad (2.2.1.1)$$

solutions (counting the multiplicity).

Proof:

See Garcia and Li (on the number of solution to polynomial system of equation, Theorem 3.1, p. 543, [19]). Their proof based on homotopy approaches. Noether and Van der Waerden [37], and Friedland [17] gave a proof using classical Bavout's Theorem [43].

As a result of this theorem we prove the following Lemma.

Lemma (2.2.2)

i. The system (2.1.3) has $n!$ solutions.

ii. If $z = (z_1, z_2, \ldots, z_n)$ is a solution of $P(z) = S$. Then for any permutation μ of $(1,2,\ldots,n)$, $z = (z_{\mu(1)}, z_{\mu(2)}, \ldots, z_{\mu(n)})$ is also a solution.

<u>Proof</u>:

i. A straightforward consequence of the Newton's identities implies that 0 is the only solution of $P(z) = 0$. Therefore by Theorem (2.2.1), $P(z) = S$ has $n!$ solutions.

ii. It is a trivial consequence of symmetry of $P(z) = S$.

Let us consider the homotopy $H : (0,1) \times U \times C^n \to C^n$ defined by

$$H(t,a,z) = (1-t)P(a) - P(z) + ts \qquad (2.2.2)$$

where

$$U = \{(z_1, z_2, \ldots, z_n) \in C^n : z_i \neq z_j \quad \text{for} \quad i \neq j\}$$

is an open subset of C^n.

<u>Lemma (2.2.3)</u>

For almost every $a \in U$ (in the sense of Bair category or Lebesgue measure) 0 is a regular value of $H_a(t,z)$, where

$$H_a : (0,1) \times C^n \to C^n$$

is defined by

$$H_a(t,z) = H(t,a,z) .$$

Proof:

Let $H(t,a,z) = 0$ for some $(t,a,z) \in (0,1) \times U \times C^n$.

Then

$$D_a H(t,a,z) = (1-t)p'(a)\cdot \tag{2.2.3}$$

Therefore

$$\det D_a H(t,a,z) = (1-t)\det p'(a)$$

$$= (1-t)\det \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 2a_1 & 2a_2 & \cdots & 2a_n \\ \vdots & & & \vdots \\ na_1^{n-1} & na_2^{n-1} & \cdots & na_n^{n-1} \end{vmatrix}$$

$$= n!(1-t)\det \begin{vmatrix} 1 & 1 & \cdots & 1 \\ a_1 & a_2 & \cdots & a_n \\ \vdots & & & \\ a_1^{n-1} & a_2^{n-1} & \cdots & a_n^{n-1} \end{vmatrix} .$$

This involves the determinant of the Vandermond matrix, which is well-known, and we obtain

$$\det D_a H(t,a,z) = (1-t)n! \prod_{1\le i<j\le n} (a_j - a_i) \neq 0$$

$$\forall\, a \in U .$$

Thus for $t \neq 1$, $D_a H(t,a,z)$ is nonsingular. Hence for $t \neq 1$

$$\text{Range } DH(t,a,z) \supseteq \text{Range } D_a H(t,a,z) = C^n . \qquad (2.2.4)$$

Therefore DH has rank n. This means 0 is a regular value of $H(t,a,z)$. Thus by the generalized Sard's theorem (1.3.2) for almost every $a \in U$, 0 is a regular value for $H_a(t,z) = H(t,a,z)$.

As we mentioned in section three of chapter one, each component of

$$H_a^{-1}(0) = \{(t,z) : H_a(t,z) = 0\}$$

is a smooth curve which is diffeomorphic to either a circle or an open interval. Let a be choosen so that 0 is a regular value of $H_a(t,z)$. Let

$$\Gamma_a = \{(t(\lambda), z(\lambda)) : 0 \leq \lambda < \lambda_1,\ (t(0), z(0)) = (0,a)\}$$

be a component of $H_a^{-1}(0)$ which contains $(0,a)$ and it is parameterized by a parameter λ. Differentiating $H_a(t(\lambda), z(\lambda)) = 0$ with respect to the parameter λ implies

$$D_z H_a . \dot{Z} + D_t H_a . \dot{t} = 0 \qquad (2.2.5)$$

where $\dot{z} = (\frac{\partial z_1}{\partial \lambda}, \frac{\partial z_2}{\partial \lambda}, \ldots, \frac{\partial z_n}{\partial \lambda})^t$ and $\dot{t} = \frac{dt}{d\lambda}$.

<u>Theorem (2.2.4)</u>

Let 0 be a regular value for H_a, then with the above notation $t(\lambda)$ is a monotonic function of λ.

Moreover $\dot{t} \neq 0$ for any $(t,z) \in \Gamma_a$.

Proof:

Let us rewrite $H_a : C^n \times [0,1] \to C^n$ and

$$z = (x_1 + iy_1, \ldots, x_n + iy_n)$$

in terms of their real and imaginary components obtaining

$$G : \mathbb{R}^m \times [0,1] \to \mathbb{R}^m$$

and $W = (x_1, y_1, \ldots, x_n, y_n)$ where $m = 2n$. By Theorem (1.4.3) we have either

$$\operatorname{sgn} \dot{t} = \operatorname{sgn} \det G'_w \text{ for all } \lambda \qquad (2.2.6)$$

or

$$\operatorname{sgn} \dot{t} = -\operatorname{sgn} \det G'_w \text{ for all } \lambda \qquad (2.2.6.1)$$

where G'_w is the $m \times m$ Jacobian of G with respect to x and y. It is shown in [21] (also in [5]) that $\det G'_w \geq 0$. Hence by (2.2.6) and (2.2.6.1), $\dot{t} \geq 0$ or $\dot{t} \leq 0$ for all λ. This proves the monotonicity of t. If $\dot{t} = 0$ for some $(t,z) \in \Gamma_a$, then (2.2.5) implies that $D_z H_a = -P'(z)$ is singular. So z is a multiple root of $H_a(t,z)$ which contradicts the result of Lemma (2.2.2). This completes the proof of the theorem.

Without loss of generality we may assume $\dot{t} > 0$. If we start at $t = 0$ and increase the parameter, then $t(\lambda)$ increases. This means t cannot turn around and return back to $t = 0$. In other words, the monotonicity of t implies that Γ_a cannot be diffeomorphic to a circle, and each component of Γ_a with $t = 0$ at one end will either go to infinity or to a solution of $P(z) = s$.

At this stage it is important to know whether the component of Γ_a will diverge to infinity or not. For this we prove the following:

Lemma (2.2.5)

The connected component Γ_a of $H_a^{-1}(0)$ which contains a is a bounded curve.

Proof:

Suppose this is not the case. Then there is a $\bar{\lambda} < 1$, such that $(t(\lambda), z(\lambda))$ diverges to infinity as λ goes to $\bar{\lambda}$. Since for any k, $P_k(z)$ is a homogeneous polynomial of degree k (i.e. $P_k(\lambda z) = \lambda^k P_k(z)$),

$$P_k\left(\frac{z(\lambda)}{\|z(\lambda)\|}\right) = \frac{1}{\|z(\lambda)\|^k} P_k(z(\lambda)) = \frac{1}{\|z(\lambda)\|^k}[tS_k + (1-t)P_k(a)] \ .$$

So $P_k\left(\frac{z(\lambda)}{\|z(\lambda)\|}\right)$ goes to zero as λ goes to $\bar{\lambda}$ for all k. Hence, if $\bar{z}$ is a cluster point of $\frac{z(\lambda)}{\|z(\lambda)\|}$ we have, $P(\bar{z}) = 0$ with $\|\bar{z}\| = 1$. However, this cannot be true because 0 is the only solution of $P(z) = 0$.

§(2.3) Following the curve

Integrating the initial value problem

$$\begin{cases} P'(z) \cdot \dot{z} = S - P(a) \\ \qquad z(0) = a \end{cases} \qquad (2.3.1)$$

is what we mean by following the curve. Since $P'(z)$ is nonsingular (2.3.1) can be written as

$$\begin{cases} \dfrac{dz}{d\lambda} = (P'(z))^{-1} \, (S - P(a)) \\ z(0) = a \, . \end{cases} \tag{2.3.2}$$

Nowadays, there are efficient initial value problem solvers which can be used to solve this problem. In general we form the Jacobian matrix $p'(z)$ and solve for $\dot{z}$ which satisfies in (2.3.2). However this procedure uses $O(n^3)$ operations each time. Also $O(n^2)$ function evaluation are needed for computation of $p'(z)$. Reducing the cost of computation is the purpose of this section. We shall utilize the symmetrical structure of $p(z)$ in such a way that $(p'(z))^{-1}$ may be obtained by $O(n^2)$ operations. To do this we introduce the following:

Suppose we have a function g which is known at a set of disjoint points $\{x_1, x_2, \ldots, x_n\}$. Let us define

$$q_j(x) = \prod_{\substack{i=1 \\ i \neq j}}^{n} \frac{(x - x_i)}{(x_j - x_i)} \, , \tag{2.3.3}$$

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

Clearly $q_j(x_i) = \delta_{ij}$ (the so called Kronecker delta). The interpolating polynomial defined by

$$y(x) = \sum_{j=1}^{n} q_j(x) g(x_j) \tag{2.3.4}$$

is called the Lagrangian interpolating polynomial. It is well-known that the polynomials $q_1, q_2, \ldots, q_n$ defined by (2.3.3) are linearly independent and form a basis of vector space V of dimension n.

Now let us define the following notation which will be needed in the next proposition. For each $i = 1,2, \ldots n$ define $\sigma_1^i, \ldots, \sigma_{n-1}^i$ as in (2.1.5) with z_i deleted. For example

$$\begin{aligned}
\sigma_1^1 &= -(z_2 + z_3 + \cdots + z_n) \\
\sigma_2^1 &= (z_2 z_3 + \cdots + z_{n-1} z_n) \\
&\;\vdots \\
\sigma_{n-1}^1 &= (-1)^{n-1} z_2 z_3 \cdots z_n .
\end{aligned} \tag{2.3.4.1}$$

Also we define

$$\sigma_0^j = 1 \qquad i = 1,2,\ldots,n .$$

<u>Proposition (2.3.1)</u>

The differential equations (2.3.2) can be written explicitly in the following form:

$$\begin{cases} \dfrac{dz_i}{dt} = \dfrac{(-1)^{n+1}}{\prod\limits_{\substack{k=1 \\ k\neq i}}^{n} (z_k - z_i)} \cdot \sum\limits_{j=1}^{n} \dfrac{\sigma_{n-j}^i}{j} (S_j - P_j(a)) \\ z_i(0) = a_i \qquad\qquad i = 1,2,\ldots,n . \end{cases} \tag{2.3.5}$$

Proof:

We make use of the Lagrangian interpolation formula

$$g_j(z) = \sum_{i=1}^{n} g_j(z_i) q_i(z) \tag{2.3.6}$$

for functions $g_j(z)$ defined by

$$g_j(z) = z^j , \qquad j = 0,1,\ldots,n-1 , \tag{2.3.7}$$

and $q_i(z)$ defined by

$$q_i(z) = \prod_{\substack{j=1 \\ j \neq i}}^{n} \frac{(z - z_j)}{(z_i - z_j)} , \; i = 1,2,\ldots,n . \tag{2.3.7.1}$$

In other words we have

$$z^j = \sum_{i=1}^{n} z_i^j \, q_i(z), \qquad j = 0,1,\ldots,n-1 . \tag{2.3.8}$$

Let us rewrite (2.3.8) in the matrix form

$$\begin{pmatrix} 1 \\ z \\ \cdot \\ \cdot \\ \cdot \\ z^{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ z_1 & z_2 & \cdots & z_n \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ z_1^{n-1} & z_2^{n-1} & \cdots & z_n^{n-1} \end{pmatrix} \begin{pmatrix} q_1(z) \\ q_2(z) \\ \cdot \\ \cdot \\ \cdot \\ q_n(z) \end{pmatrix} .$$

Multiply both sides of this equation by

$$\begin{pmatrix} 1 & & & & \\ & 2 & & 0 & \\ & & 3 & & \\ & 0 & & \ddots & \\ & & & & n \end{pmatrix}$$

we thus get the following equation,

$$Z^* = WQ \tag{2.3.9}$$

where

$$W = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2z_1 & 2z_2 & & 2z_n \\ \vdots & \vdots & & \vdots \\ nz_1^{n-1} & nz_2^{n-1} & & nz_n^{n-1} \end{pmatrix}$$

$$Q(z) = (q_1(z), q_2(z), \ldots, q_n(z))^t$$

and

$$Z^* = (1, 2Z, 3Z^2, \ldots, nz^{n-1}) \quad .$$

Let

$$W^{-1} = \begin{pmatrix} a_{11} & \cdots\cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots\cdots & a_{nn} \end{pmatrix}$$

we can rewrite (2.3.9) as

$$\begin{pmatrix} q_1(z) \\ q_2(z) \\ \cdot \\ \cdot \\ \cdot \\ q_n(z) \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} 1 \\ 2z \\ \cdot \\ \cdot \\ \cdot \\ nz^{n-1} \end{pmatrix} \cdot \qquad (2.3.10)$$

Also the polynomials (2.3.7.1) can be written as the following

$$q_i(z) = \frac{(-1)^{n+1}}{\prod\limits_{\substack{k=1 \\ k\neq i}}^{n} (z_k - z_i)} [\sigma^i_{n-1} + \sigma^i_{n-2} z + \cdots + \sigma^i_0 \, z^{n-1}], \qquad (2.3.11)$$

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

Comparing the coefficients on both sides of (2.3.10), we get

$$a_{ij} = \frac{(-1)^{n+1}}{\prod\limits_{\substack{k=1 \\ k\neq i}}^{n} (z_k - z_i)} \cdot \frac{1}{j} \sigma^i_{n-j} \quad . \qquad (2.3.12)$$

The differential equations (2.3.2) thus become as in (2.3.5).

Based on the recursion formula

$$\sigma^i_j = \sigma_j + z_i \, \sigma^i_{j-1} \qquad (2.3.13)$$

for $1 \leq i \leq n$, $1 \leq j \leq n-1$, The computation of σ^i_{n-j} involves only $O(n^2)$ operations. The values σ_j's may be obtained by (2.1.6), using forward substition which needs

$\frac{n(n-1)}{2}$ operations, while the total operations needed for the second term in (2.3.13) is $(n-1)^2$.

§(2.4) Generalization

In this section we will discuss the case that two zeros of $f(z)$ are equal (multiple roots), or nearly equal. Without loss of generality we may assume that z_1 and z_2 are such zeros. In this case equations (2.3.3) are not suitable to us, because evaluation of $\frac{dz_1}{d\lambda}$ and $\frac{dz_2}{d\lambda}$ is too expensive. Because

$$\frac{1}{\prod_{k\neq i}(z_k - z_i)}$$

tends to infinity as λ tends to 1 for $i = 1,2$. To avoid this difficulty, we write a new differential equations whose solutions are easily led to the solutions of $f(z)$. We define

$$A_j = \frac{\sigma^1_{n-j}}{\prod_{k\neq 1}(z_k - z_1)} + \frac{\sigma^2_{n-j}}{\prod_{k\neq 2}(z_k - z_2)} \qquad (2.4.1)$$

$$j = 1,2,\ldots,n$$

Proposition (2.4.1)

For $1 \leq j \leq n$, $A_j(z)$ is a rational function of z and contains no factor of the form $(z_2 - z_1)^{-1}$.

Proof:

Let us write $A_j(z)$ as follows

$$A_j(z) = \frac{1}{\prod_{k\neq 1,2}^{n} (z_k-z_1)(z_k-z_2)} \times \frac{B_j(z)}{(z_2-z_1)} \qquad (2.4.2)$$

where

$$B_j(z) = \sigma^1_{n-j} \prod_{k\neq 1,2}^{n} (z_k-z_2) - \sigma^2_{n-j} \prod_{k\neq 1,2}^{n} (z_k-z_1) . \qquad (2.4.3)$$

For $i \neq j$ let us define

$$S^i_j(\sigma^i_k) \qquad i,j,k = 1,2,\ldots,n$$

to be the same as σ^i_k with z_j changed to z_i. With these notations we have

$$S^1_2(\sigma^1_{n-j}) = \sigma^2_{n-j} \qquad (2.4.4)$$

$$S^2_1(\sigma^2_{n-j}) = \sigma^1_{n-j} . \qquad (2.4.4.1)$$

Let us view $B_j(z)$ as a function of z_1, and evaluate this function at z_2. Then because of equations (2.4.4) and (2.4.4.1) we have $B_j(z_2) = 0$. That is

$$B_j(z) = (z_2 - z_1) q_j(z) \ , \qquad (2.4.5)$$

For some polynomial $q_j(z)$. Thus

$$\lim_{z_2\to z_1} \frac{B_j(z_2)-B_j(z_1)}{z_2 - z_1} = \frac{\partial}{\partial z_1} B_j \qquad (2.4.6)$$

$$= -\frac{\partial}{\partial z_1} \sigma^2_{n-j} \prod_{k\neq 1,2}^{n} (z_k - z_1)$$

A repeat use of recursion formula (2.3.13) combined with equation (2.4.6) imply

$$q_j(z) = -\frac{\partial}{\partial z_1} \sum_{m=0}^{n-j} \sigma_{n-j-m} \cdot z_2^m \prod_{k\neq 1,2}^{n} (z_k - z_1)$$

$$= -\sum_{m=0}^{n-j} z_2^m \frac{\partial}{\partial z_1} \left(\sigma_{n-j-m} \prod_{k\neq 1,2}^{n} (z_k - z_1)\right). \quad (2.4.7)$$

Hence

$$A_j(z) = q_j(z) \prod_{k\neq 1,2}^{n} \frac{1}{(z_k - z_1)(z_k - z_2)}, \quad j = 1,2,\ldots,n .$$

Proposition (2.4.2)

For $1 \leq j \leq n$, let $A_j^*(z)$ and $g_j^*(z)$ be defined by

$$A_j^*(z) = \frac{z_2 \sigma_{n-j}^1}{\prod_{k\neq 1}^{n} (z_k - z_1)} + \frac{z_1 \sigma_{n-j}^2}{\prod_{k\neq 2}^{n} (z_k - z_2)} \quad (2.4.8)$$

$$q_j^*(z) = -\frac{\partial}{\partial z_1} z_1 \sigma_{n-j}^2 \prod_{k\neq 1,2}^{n} (z_k - z_1) . \quad (2.4.8.1)$$

Then

$$A_j^*(z) = \frac{q_j^*(z)}{\prod_{k\neq 1,2}^{n} (z_k - z_1)(z_k - z_2)} . \quad (2.4.9)$$

The proof of this proposition is similar to the proof of proposition (2.4.2) and hence is left out.

Let us define

$$v_1 = (A_1(z), A_2(z), \ldots, A_n(z))^t , \qquad (2.4.10)$$

$$v_2 = (A_1^*(z), A_2^*(z), \ldots, A_n^*(z))^t . \qquad (2.4.10.1)$$

Then the initial value problem (2.3.1) can be written as

$$\frac{dz_i}{d\lambda} = (-1)^{n+1} v_i^t \cdot (s - P(a)) \qquad i = 1,2$$

$$\frac{dz_j}{d\lambda} = \frac{(-1)^{n+1}}{\prod\limits_{k\neq j}^{n} (z_k - z_j)} \sum_{\ell=1}^{n} \frac{1}{\ell} \sigma_{n-\ell}^{j} (s_\ell - P_\ell(a)) \qquad (2.4.11)$$

$$j = 3, \ldots, n$$

with the initial value $z(0) = (a_1 + a_2, a_1 a_2, a_3, \ldots, a_n)^t$.

§(2.5) Numerical results

In this section a series of numerical examples are given. We have used the ordinary differential equations solver subroutine "DE" (see "Computer solution of ordinary differential equations: The initial value problem" by L.F. Shampine and M.K. Gordon). Subroutine "DE" is probably the most sophisticated and advanced subroutine of this kind.

Briefly subroutine "DE" integrates a system of up to 20 first order ordinary differential equations of the form

$$\frac{dy_i}{dt} = F(t, y_1, y_2, \ldots, y_n), \quad i = 1, \ldots, n \; . \tag{2.5.1}$$

The complete description of this subroutine can be found in [40] (the computer program of "DE" is given on pp. 186-209). The following procedure employing the subroutine "DE" was used.

i) The computer's random function routine generates the starting vectors $z_1, z_2, \ldots, z_n$. We choose two small positive numbers ε_1, ε_2 to list the absolute and relative errors respectively.

ii) Designate the exist time, called tout, and call the subroutine "DE" in order to follow the path.

iii) We write down the output of (ii) and compare them.

a) If all solutions are isolated in the sense that they are far from each other by at least $\varepsilon > 0$, we use equation (2.3.5) to initialize the differential equation and go to (iv).

b) Otherwise we use equations (2.4.11) to initialize the differential equations and go to (iv).

iv) If tout = 1 we read the solutions and stop.

v) Otherwise go to (ii).

In the following tables we have shown some numerical results for four systems of equations with known solutions in order to evaluate the accuracy of our algorithm.

n	Starting Vectors
1	(.580113649E + 00, .950512735E + 00)
2	(.786371425E + 00, .297620264E + 00)
3	(.453699900E + 00, .626194160E - 02)
4	(.275736426E + 00, .305650944E + 00)
5	(.689100711E + 00, .382662239E + 00)
6	(.132902705E + 00, .831857903E + 00)
7	(.582979796E + 00, .986253383E - 01)
8	(.276548455E + 00, .620446028E + 00)
9	(.835029668E - 01, .990377121E + 00)
10	(.979346943E + 00, .693884438E + 00)
11	(.934477014E + 00, .212092955E + 00)
12	(.130652748E + 00, .862596980E + 00)
13	(.818909294E + 00, .862596980E + 00)
14	(.187994091E - 01, .314116017E + 00)
15	(.765182100E + 00, .941526108E + 00)

Table (2.1)

The starting vectors for various problems came from Table (2.1). Its first 5 vectors have been used as the starting vector for examples 1 and 4. Also its first 10 vectors are the starting vectors for example 2. Finally all these vectors are used as the starting vectors for example 3.

Example 1. We solve system (2.1.3) for $n = 5$. $s_1,\ldots,s_5$ are chosen so that the exact solutions are 5,9,7,4,2. The computed results are as follows:

n	input vector	constant vector
1	.27000E + 01	(.242150E + 02, -.194271E + 01)
2	.17500E + 03	(.147520E + 03, -.227252E + 01)
3	.12690E + 04	(.127004E + 04 -.116076E + 01)
4	.98590E + 04	(.985999E + 04, -.411588E + 00)
5	.80037E + 05	(.800367E + 05, .100286E + 01)

Table (2.2)

Homotopy solutions at time t = 1
z(1) = (.499999999E + 01, -.644220655E - 08)
z(2) = (.900000000E + 01, -.194643929E - 09)
z(3) = (.699999999E + 01, .119773958E - 09)
z(4) = (.400000000E + 01, .627337758E - 08)
z(5) = (.200000000E + 01, -.837647611E - 09)

Table (2.2.1)

The relative and absolute errors in this problem are respectively $\epsilon_1 = 10^{-15}, \epsilon_2 = 10^{-12}$.

<u>Example 2</u>. This problem solves system (2.1.3) for $n = 10$. $s_1,\ldots,s_{10}$ are chosen so that the exact solutions are $1,2,3,\ldots,10$, (in same order). Here $\epsilon_1 = 10^{-12}$, $\epsilon_2 = 10^{-10}$. The computed results are as follows:

n	input vector	constant vector
1	.5500000E + 02	(.501597E + 02, -.517790E + 01)
2	.3850000E + 03	(.385669E + 03, -.447630E + 01)
3	.3025000E + 04	(.302715E + 04, -.134409E + 01)
4	.2533300E + 05	(.253342E + 05, -.130606E + 00)
5	.2208250E + 06	(.220826E + 06, -.309642E + 00)
6	.1978405E + 07	(.197841E + 07, .695913E + 00)
7	.1808043E + 08	(.180804E + 08, .242440E + 01)
8	.1677313E + 09	(.167731E + 09, .283408E + 01)
9	.1574305E + 10	(.157440E + 10, .233227E + 01)
10	.1491434E + 11	(.149143E + 11, .214617E + 01)

Table (2.3)

Homotopy solutions at time t = 1
z(1) = (.500000004E + 01, -.157844480E - 06)
z(2) = (.900000000E + 01, -.115510367E - 07)
z(3) = (.599999998E + 01, .146901334E - 06)
z(4) = (.399999996E + 01, .119657234E - 06)
z(5) = (.999999999E + 01, .116296633E - 08)
z(6) = (.300000003E + 01, -.544245438E - 07)
z(7) = (.799999997E + 01, .708308777E - 07)
z(8) = (.199999999E + 01, .136081468E - 07)
z(9) = (.100000000E + 01, -.125990729E - 08)
z(10) = (.700000003E + 01, -.128213396E - 06)

Table (2.3.1)

Example 3. Here we solve system (2.1.3) for $n = 15$. $s_1, s_2, \ldots, s_{15}$ are chosen so that the exact solutions are $1, 2, \ldots, 15$, (in some order). We choose $\varepsilon_1 = \varepsilon_2 = 10^{-8}$. The computed results are as follows:

n	input vector	constant vector
1	.1200000E + 03	(.112498E + 03, -.804895E + 01)
2	.1240000E + 04	(.124059E + 04, -.743638E + 01)
3	.1440000E + 05	(.144035E + 05, -.301077E + 01)
4	.1783120E + 06	(.178315E + 06, -.259791E + 00)
5	.2299200E + 07	(.229920E + 07, .923576E + 00)
6	.3048292E + 08	(.304829E + 08, .251047E + 01)
7	.4124208E + 09	(.412421E + 09, .283309E + 01)
8	.6666482E + 10	(.566648E + 10, -.132393E + 00)
9	.7880094E + 11	(.788009E + 11, -.326317E + 01)
10	.1166533E + 13	(.119653E + 13, -.162573E + 01)
11	.1866217E + 14	(.156622E + 14, .262221E + 01)
12	.8231603E + 15	(.223160E + 15, .365233E + 01)
13	.3197504E + 16	(.319750E + 16, -.390705E + 01)
14	.4663402E + 17	(.460340E + 17, -.134212E + 02)
15	.6664785E + 18	(.665479E + 18, -.172283E + 02)

Table (2.4)

Homotopy solutions at time t = 1
z(1) = (.900039039E + 01, .177185166E - 02)
z(2) = (.119999114E + 02, -.182960120E - 03)
z(3) = (.799988123E + 01, -.218749510E - 02)
z(4) = (.600017487E + 01, -.144389314E - 02)
z(5) = (.700006205E + 01, .205918686E - 02)
z(6) = (.999997814E + 00, -.165688702E - 06)
z(7) = (.999971337E + 01, -.110689513E - 02)
z(8) = (.200001713E + 01, -.417415352E - 05)
z(9) = (.499975294E + 01, .735860613E - 03)
z(10) = (.130000293E + 02, .475078047E - 04)
z(11) = (.139999773E + 02, -.391695467E - 05)
z(12) = (.299992826E + 01, .531830915E - 04)
z(13) = (.150000010E + 02, .363097092E - 06)
z(14) = (.400017191E + 01, -.256633723E - 03)
z(15) = (.110001766E + 00, .522092661E - 03)

Table (2.4.1)

Remark. The above examples along with a lot of other numerical experiments suggest that as long as n, the number of the equations, increases, we would have less accurate results. This is due to the subroutine "DE" and presumably because this subroutine is coded for at most 21 equations. It should be mentioned that in each problem we have converted n complex variables into $2n$ real variables and with the parameter time t totally we worked with $2n+1$ real variables. Other subroutines such as "DeVerk", "D Gear" from IMSL are also used to solve these problems. But overall our numerical experiments show that among all available differential equations solvers "DE" is the best one, even though it is not very accurate for problems with relatively large dimension.

Example 4. This problem solves system (2.1.3) for $n = 5$, and its exact solutions are 2, 5, 5, 7, 8. Here we follow the curves until two of them get close to each other, that is, the distance between them be less than $\epsilon = 10^{-5}$. Then we follow the curves by integrating of differential equations (2.4.11). The relative and absolute errors in this problem are respectively $\epsilon_1 = 10^{-15}$, $\epsilon_2 = 10^{-12}$.

n	input vector	constant vector
1	.27000E + 02	(.242150E + 02, -.194271E + 01)
2	.16700E + 03	(.166520E + 03, -.227252E + 01)
3	.11130E + 04	(.111404E + 04, -.116076E + 01)
4	.77630E + 04	(.776399E + 04, .411588E + 00)
5	.55857E + 05	(.558567E + 05, .100286E + 01)

Table (2.5)

n	Homotopy solution at the time t = .899 where the distance between two curves is less than ϵ.
1	(.500000358E + 01, .226529962E - 04)
2	(.799999999E + 01, .459513577E - 10)
3	(.700000000E + 01, -.134110233E - 09)
4	(.499999642E + 01, -.226529034E - 04)
5	(.199999999E + 01, -.834203779E - 11)

Table (2.5.1)

Homotopy solutions at t = 1
z(1) = (.999999999E + 01, .220664352E - 09)
z(2) = (.799999999E + 01, .728877445E - 10)
z(3) = (.700000000E + 01, -.198322803E - 09)
z(4) = (.250000000E + 02, .101079963E - 08)
z(5) = (.199999999E + 01, -.967861485E - 10)

Table (2.5.2)

CHAPTER THREE

REAL ZEROS OF A FUNCTION

§(3.1) Introduction

In this chapter, we shall be finding the real roots of

$$f(z) = 0 \tag{3.1.1}$$

in the interval $[-R,R]$, $(R > 0)$, where $f(z)$ is an analytic function on an open set U containing the interval $[-R,R]$. We shall use Cauchy's formula

$$S_0 = \frac{1}{2\pi i} \int_{\partial S} \frac{f'(z)}{f(z)} \, dz \tag{3.1.2}$$

where S is a rectangle with vertices $\pm R \pm i\varepsilon$ and entirely contained in U and ∂S is the boundary of S, (ε is a small positive number). S_0 is the number of zeros of $f(z)$ inside the contour ∂S, and hence is an integer number. Thus we need only a crude estimate of S_0. When we find an approximation for S_0, say s_0, then the dimension of our problem will be the integer nearest to s_0. Having the dimension (number of zeros) we will employ the generalized Cauchy's formula

$$S_i = \frac{1}{2\pi i} \int_{\partial S} z^i \frac{f'(z)}{f(z)}\, dz \qquad i = 1,2,\ldots,n \tag{3.1.3}$$

to approximate the numbers $S_1, S_2, \ldots, S_n$, where

$$S_i = \sum_{j=1}^{n} z_j^i, \qquad i = 1,2,\ldots,n .$$

These numbers along with the homotopy continuation method presented in chapter two will determine an approximation for the real zeros of $f(z)$. In order to obtain a more accurate result, the approximated zeros $z_1, z_2, \ldots, z_n$ could be a good starting point for using the Newton's iteration method. That is to refine the approximations obtained by the homotopy continuation method, one or two-step method can be used as the final stage in calculation of the real zeros of $f(z)$.

§(3.2) Integration around the boundary of a strip

Suppose $g \in C[0,1]$ and define

$$\begin{aligned} R^{[m,1]}(g) &= \frac{1}{m} \sum_{j=0}^{m}{}' g(j/m) \\ &= \frac{1}{m}\left(\frac{1}{2}\, g(0) + g\left(\frac{1}{m}\right) + \cdots + g\left(\frac{m-1}{m}\right) + \frac{1}{2}\, g(1)\right). \end{aligned} \tag{3.2.1}$$

The prime on the summation indicates that the first and the last terms are assigned a weighting factors $\frac{1}{2}$. Using Romberg's method of integration, $R^{[m,1]}(g)$ is evaluated for mesh $m = 1,2,4,8,\ldots$.

Let us define

$$T_0^k = R^{[2^k,1]} \qquad k = 0,1,2,3,\ldots . \qquad (3.2.2)$$

Then we evaluate T_0^0, T_0^1, T_0^2, ..., and with the aid of the following recursion formula

$$T_m^k = \frac{1}{4^m-1}\left(4^m\, T_{m-1}^{k+1} - T_{m-1}^k\right) \qquad (m > 0) \qquad (3.2.3)$$

we form the following T-table:

$$\begin{array}{ccccc}
T_0^0 & & & & \\
T_0^1 & T_1^0 & & & \\
T_0^2 & T_1^1 & T_2^0 & & \\
T_0^3 & T_1^2 & T_2^1 & T_3^0 & \\
T_0^4 & T_1^3 & T_2^2 & T_3^1 & T_4^0 \\
\vdots & & & & \ddots
\end{array} \qquad (3.2.4)$$

Equation (3.2.3) implies that T_m^k is a linear combination of T_0^k, T_0^{k+1}, T_0^{k+2}, ..., T_0^{k+m}. That is

$$T_m^k = \sum_{i=0}^{m} C_m^{m-i}\, T_0^{k+i} . \qquad (3.2.5)$$

Since g is Riemann integrable on $[0,1]$, as k goes to infinity the first column of the T-table, namely the sequence $\{T_0^k\}$ converges to $\int_0^1 g(t)\,dt$.

The leading term in the error of T_m^k is of the order $(\frac{1}{2^{m+1}})^{2k+2}$, (we will see this in the next section) while that of T_0^k is $(\frac{1}{2^{k+1}})^2$. Thus T_m^k converges much more rapidly than T_0^k. After all the convergence of the first column of the T-table implies the convergence of all further columns and all diagonal sequences (T_m^k with k constant) to the same limit. We observe that dividing h successively by 2 is a sound one, since each application of the composite rule (namely, equation (3.2.1)) utilizes all of the function values previously used for smaller values of k and $m = 0$. A division of h by some other number would require introducing a complete set of new function values. Hence the above form of the Romberg's method of integration based on trapezoidal rule requires fewer function values than might at first appear. Moreover the recursion formula (3.2.4) is a simple manipulation and requires minimal computer time. The most time consuming part of the algorithm is the calculation of T_0^k, $k = 1,2,\ldots$, the original composite trapezoidal rule.

§(3.3) Error analysis

Here we consider the error estimate of numerical approximation to

$$S_i = \frac{1}{2\pi i}\int_{\partial S} z^i \frac{f'(z)}{f(z)}\,dz, \quad i = 1,2,\ldots,n \qquad (3.3.1)$$

where S denotes a narrow strip with vertices $\pm R \pm i\varepsilon$, (ε is a small positive number).

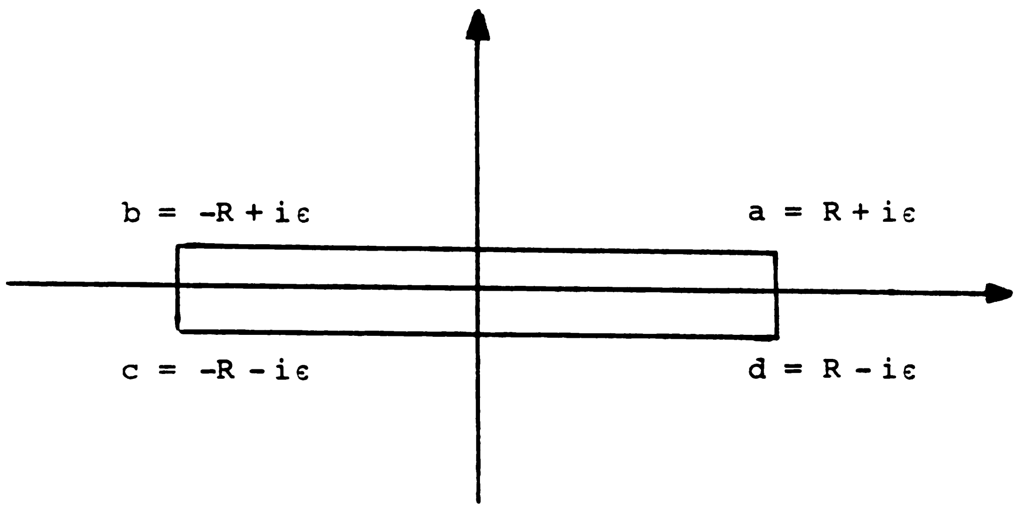


Figure (3.1)

Let us assume $M = \sup\limits_{z \in \partial S} \frac{f'(z)}{f(z)}$, and choose ε small enough (with $\varepsilon < R$). Then

$$\left|\frac{1}{2\pi i} \int_{\overline{bc}} z^i \frac{f'(z)}{f(z)}\, dz + \frac{1}{2\pi i} \int_{\overline{da}} z^i \frac{f'(z)}{f(z)}\, dz\right| \leq \frac{2\varepsilon M}{\pi}(\sqrt{2}\, R)^i \tag{3.3.2}$$

$$i = 0,1,2,\ldots \quad .$$

Having ε small enough, the amount of integration on the two vertical sides of the strip S is contributing a small amount in the whole process of integration. Let us denote this amount by the initial error E_I. Next we consider the integration along the other two sides. By using a simple change of variable we may consider the integration on the interval $[0,1]$. In order to analyze the corresponding error we define the following.

Definition (3.3.1)

A function φ is of class E (even) if

i) $\varphi(x)$ is of definite sign for $0 < x < 1$,

ii) $\varphi'(x)$ is of definite sign for $0 < x < 1/2$,

iii) $\varphi(0) = \varphi(1)$, $\varphi(x) = \varphi(1-x)$ and $\varphi(x) = \varphi(1+x)$.

Definition (3.3.2)

A function ψ is of class O (odd) if

i) $\psi(x)$ is of definite sign for $0 < x < 1$,

ii) $\psi(0) = \psi(1/2) = \psi(1)$,

iii) $\psi(x) = -\psi(1-x)$, and $\psi(x) = \psi(1+x)$.

For example $\varphi(x) = |\sin \pi x|$ and $\psi(x) = \sin 2\pi x$ are respectively of class E and O.

The following lemma is a straightforward result of the definitions.

Lemma (3.3.3)

a) If $\varphi(x)$ is of class E, then

$$\psi_1(x) = \int_0^x (\varphi(2x) - \varphi(x))\,dx ,$$

and

$$\psi_2(x) = \int_0^x \varphi(x)\,dx - x \int_0^1 \varphi(x)\,dx$$

are of class O.

b) If $\psi(x)$ is of class O, then

$$\varphi(x) = \int_0^x \psi(x)\,dx$$

is of class E.

Let us define the following functions inductively:

$$K_2(x) = \begin{cases} \frac{1}{2}\,x(1-x) & \text{if} \quad 0 \le x \le 1 \\ \\ K_2(x-n) & \text{if} \quad n \le x \le n+1 . \end{cases} \tag{3.3.3}$$

Suppose K_{2m} is defined, then we define

$$K_{2m+1}(x) = \int_0^x \overline{K}_{2m}(x)\,dx \tag{3.3.3.1}$$

where

$$\overline{K}_{2m}(x) = \frac{1}{4^m-1}\,(K_{2m}(2x) - K_{2m}(x)) . \tag{3.3.3.2}$$

Also we define

$$K_{2m+2}(x) = \int_0^x K_{2m+1}(x)\,dx . \tag{3.3.3.3}$$

It is not hard to show that the functions $K_{2m}(x)$ are all positive in the interval $[0,1]$; more exactly monotically increasing from $x = 0$ to $x = 1/2$ and from there again decreasing to zero [33].

Theorem (3.3.4)

Let $g \in C^{2m+2}[0,1]$. Then

$$T_m^k - Ig = 4^{-(m+1)k} \int_0^1 K_{2m+2}(2^k x) g^{(2m+2)}(x)\,dx \ , \qquad (3.3.4)$$

where $Ig = \int_0^1 g(x)\,dx$.

Proof:

The remainder term for trapezoidal rule is given by ([3])

$$T_0^0 - Ig = \int_0^1 \frac{t}{2}(1-t) g''(t)\,dt \ ,$$

and likewise

$$T_0^k - Ig = 4^{-k} \int_0^1 K_2(2^k t) g''(t)\,dt \ .$$

We prove the theorem by induction with respect to m. Assume

$$T_{m-1}^k - Ig = 4^{-km} \int_0^1 K_{2m}(2^k t) g^{(2m)}(t)\,dt \ , \qquad (3.3.5)$$

where K_{2m} is a function of class E. This assumption is certainly true for $m = 1$, with $K_2(x)$. Substitution of (3.2.3) into (3.3.5) gives

$$T_m^k - Ig$$
$$= \frac{1}{4^m - 1} \int_0^1 [4^m \cdot 4^{-(k+1)m} K_{2m}(2^{k+1}t) - 4^{-km} K_{2m}(2^k t)]$$
$$g^{(2m)}(t)\,dt \ .$$

Hence

$$T_m^k - Ig = \frac{4^{-km}}{4^m-1} \int_0^1 [K_{2m}(2^{k+1}t) - K_{2m}(2^k t)] g^{(2m)}(x)\,dx. \quad (3.3.6)$$

By the way that kernel functions are defined and lemma (3.3.3) $\bar{K}_{2m}(x)$ is of class E. Also

$$K_{2m+1}(1) = \int_0^1 \bar{K}_{2m}(x)\,dx = \frac{1}{4^m-1} \int_0^1 (K_{2m}(2x) - K_{2m}(x))\,dx$$

$$= \frac{1}{4^m-1} \left[\frac{1}{2} \int_0^2 K_{2m}(x)\,dx - \int_0^1 K_{2m}(x)\,dx\right] = 0 . \quad (3.3.7)$$

Similarly

$$K_{2m+2}(1) = \int_0^1 K_{2m+1}(x)\,dx = \int_{-1/2}^{1/2} K_{2m+1}(x)\,dx = 0 . \quad (3.3.8)$$

Clearly $\bar{K}_{2m}(x)$ is also even and has period one. According to this and equations (3.3.7) and (3.3.8) a twofold integration by parts yields

$$\int_0^1 \bar{K}_{2m}(2^k t) g^{(2m)}(t)\,dt = 4^{-k} \int_0^1 K_{2m+2}(2^k t) g^{(2m+2)}(t)\,dt .$$

Therefore

$$T_m^k - Ig = 4^{-(m+1)k} \int_0^1 K_{2m+2}(2^k t) g^{(2m+2)}(t)\,dt .$$

This completes the proof of the theorem.

Let us define

$$E_{m,k} = T_m^k - \int_0^1 g(t)\,dt$$

Then equation (3.3.4) implies

$$|E_{m,k}| \leq 4^{-(m+1)k} \left|\int_0^1 K_{2m+2}(2^k t) g^{(2m+2)}(t)\,dt\right|$$

$$= 4^{-(m+1)k} |g^{(2m+2)}(\xi)| \left|\int_0^1 K_{2m+2}(2^k t)\,dt\right| \quad (3.3.9)$$

for some $\xi \in [0,1]$, (the last equation is true because the kernel function K_{2m+2} has a fixed sign in $[0,1]$). In order to arrive at an exact error estimate, we need to derive a compact form for

$$K = \int_0^1 K_{2m+2}(t)\,dt \ .$$

Let us apply equation (3.3.4) for $g(x) = \frac{x^{2m+2}}{(2m+2)!}$, then

$$T_m^0 - \int_0^1 g(t)\,dt = \int_0^1 K_{2m+2}(t)\,dt = K \ .$$

Because of the recursion formula (3.2.3), every entry of the T-table is a linear combination of values of the first column, that is

$$T_m^k = \sum_{i=0}^{m} C_m^{m-i} \; T_0^{k+1} \quad (3.3.10)$$

where C_i^j are independent of k. In fact with the assumption $C_m^{m-1} = C_m^{-1} = 0$ we have

$$C_m^i = \frac{1}{4^m - 1} (4^m C_{m-1}^i - C_{m-1}^{i-1}) \ .$$

If the Euler-MacLaurin sum for $g(x) = \frac{x^{2m+2}}{(2m+2)!}$ is written in terms of Bernolli coefficients (see [3], p. 207), we can see that

$$K = (-1)^m \frac{B_{2m+2}\, 4^{-m(m+1)}}{(2m+2)!}\, t_m\,(4^{m+1}) \ ,$$

where

$$t_m(4^{m+1}) = \frac{(1-\frac{4^{m+1}}{4})(1-\frac{4^{m+1}}{4^2}) \cdots (1-\frac{4^{m+1}}{4^m})}{(1-\frac{1}{4})(1-\frac{1}{4^2}) \cdots (1-\frac{1}{4^m})} \ ,$$

and B_{2m+2} is the $2m+2$<u>nd</u> Bernoulli number. Therefore

$$K = \frac{B_{2m+2}}{(2m+2)!\,2^{m(m+1)}} \ . \qquad (3.3.11)$$

Thus because of (3.3.9) we have

$$|E_{m,k}| \le \frac{4^{-(m+1)k} \cdot B_{2m+2}}{(2m+2)!\,2^{m(m+1)}} \, |g^{(2m+2)}_{(\xi)}| \qquad (3.3.12)$$

$$= \frac{4^{-(m+1)k}}{(2m+2)!} \, \delta \, |g^{(2m+2)}_{(\xi)}| \ ,$$

where $\delta = \frac{B_{2m+2}}{2^{m(m+1)}}$ and goes superlinearly to zero [33]. Let ρ be the shortest distance of any singularity of $g(z)$ from the interval $[0,1]$. Then for any $\delta_0 > 0$ there exists a constant such that

$$|g^{(2m+2)}_{(\xi)}| \le \frac{(2m+2)!\,K(\delta_0)}{(\rho-\delta_0)^{2m+2}} \ .$$

Let us consider the special case where $g(z)$ has the form

$$g(z) = z^i \frac{f'(z)}{f(z)} .$$

That is $g(z)$ may have only simple singularities. Thus a constant K exists such that

$$|g^{(2m+2)}_{(z)}| \leq \frac{(2m+2)!K}{\rho^{2m+2}} \tag{3.3.13}$$

In order to consider the error of integration on $\overline{ab}$ we use transformation $z = a + (b-a)t$. Thus $d = 2\rho R$ is the shortest distance of any singularity of g from line segment $\overline{ab}$.

$$\begin{aligned} |E_{m,k}| &\leq \frac{4^{-(m+1)k}}{(2m+2)!} \delta \cdot \frac{(2m+2)!K}{\rho^{2m+2}} 2R \\ &= \delta 4^{-(m+1)k} \left(\frac{2R}{d}\right)^{2m+2} K \cdot 2R \\ &= \delta 2^{(1-k)(m+1)} \cdot \frac{K}{d^{2m+2}} R^{2m+4} \end{aligned}$$

Let $E^1_{m,k}$ and $E^2_{m,k}$ be the errors corresponding to the horizontal sides $\overline{ab}$ and $\overline{cd}$ respectively. Then

$$|E^i_{m,k}| \leq \delta R^{2m+2} d^{-(2m+2)} K \cdot 2^{(1-k)(m+1)} \quad i = 1,2 . \tag{3.3.14}$$

Let $E_F = E^1_{m,k} + E^2_{m,k}$, then

$$\begin{aligned} |E_F| &\leq 2\delta R^{2m+4} d^{-(2m+2)} K \cdot 2^{(1-k)(m+1)} \\ &\leq 2\delta R^{2m+4} \cdot K\, 2^{(1-k)(m+1)} \cdot \varepsilon^{-(2m+2)} \end{aligned} \tag{3.3.15}$$

Now let us consider the corresponding error for S_0 and use the fact that S_0 is an integer number. In this case equation (3.3.2) is of the form

$$|E_I| \le \frac{2\epsilon}{\pi} M .$$

Note that the integration along the vertical sides of the strip S are ignored. In order to determine S_0 it is enough to assume the following

$$|E_I| \le \frac{1}{2^3}, \quad |E_F| \le \frac{1}{2^3} \tag{3.3.16}$$

Hence

$$2\delta R^{2m+4} K \; 2^{(1-k)(m+1)} \; \epsilon^{-(2m+2)} \le \frac{1}{2^3} \tag{3.3.17}$$

and

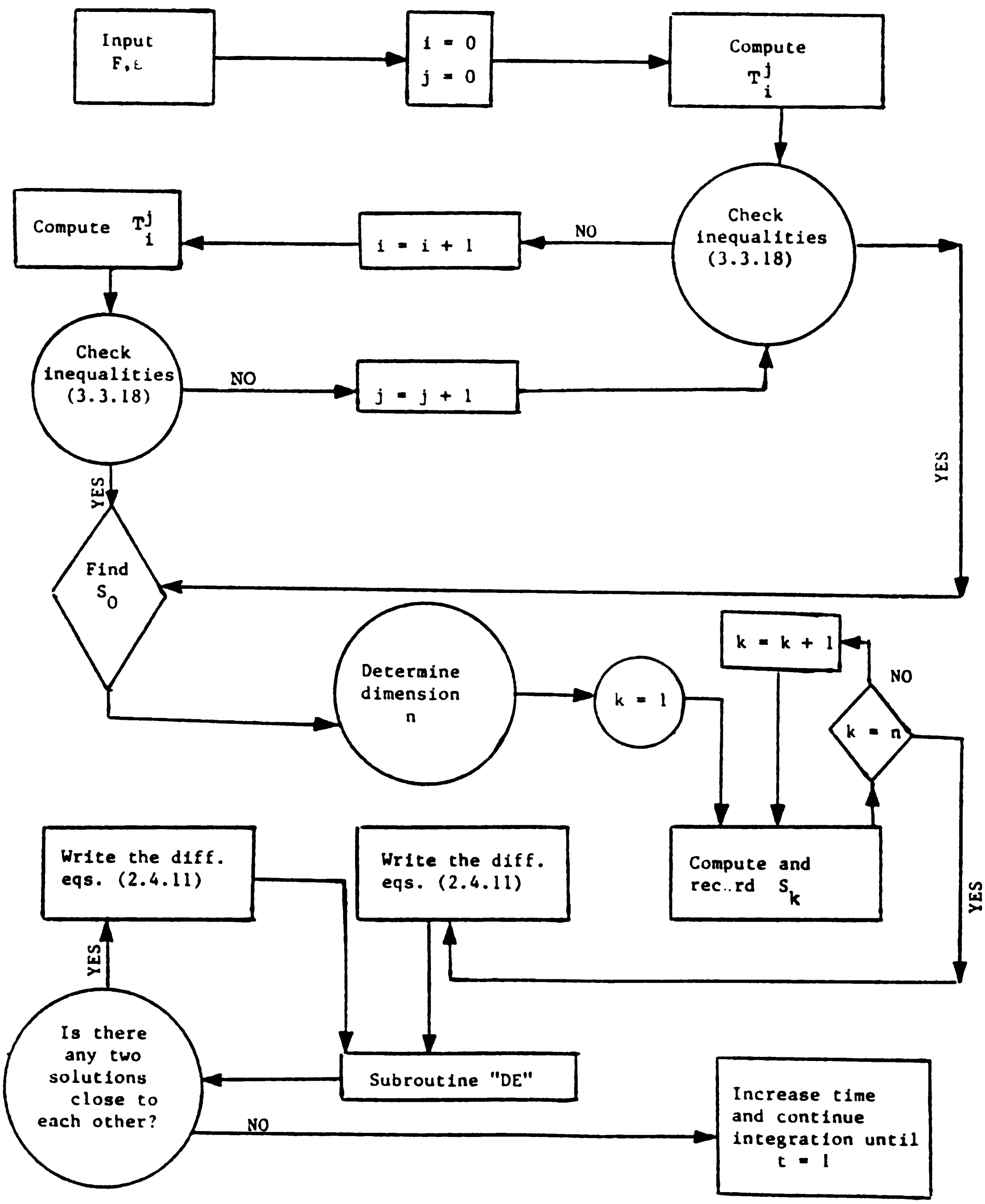
$$\frac{2\epsilon}{\pi} M \le \frac{1}{2^3} . \tag{3.3.17.1}$$

This leads to the following inequalities

$$C \; 2^{\left(\frac{1-k}{2}\right)} \le \epsilon \le \frac{\pi}{2^4 M} , \tag{3.3.18}$$

where $C = R(16\delta R^2 K)^{\frac{1}{2m+2}}$.

Based on the above argument we present the following flow chart. This flow chart gives a general framework for the algorithm.



Flow chart (3.1)

§(3.4) Numerical results

A series of computer experiments were conducted to show the accuracy of the algorithm. As the Flow chart shows, we limit our discussion here to the homotopy type algorithm of chapter two of this thesis. The computed results for the well known Bessel function are shown in Table (3.1).

Example:

Here we consider the real roots of $J_1(z)$ in the interval $[-4\pi, 4\pi]$, where $J_p(z)$ is the Bessel function of the first kind of order P defined by

$$J_p(z) = \sum_{m=0}^{\infty} \frac{(-1)^m (z/2)^{P+2m}}{m!\Gamma(m+P+1)}$$

and is the solution of the following differentiable equation

$$W'' + z^{-1}W' + (1 - P^2 z^{-2})W = 0 \ .$$

First of all by a crude estimate of S_0 we observe that $J_1(z)$ has seven zeros in $[-4\pi, 4\pi]$. In fact $z = 0$ is a zero of $J_1(z)$ (trivial solution). Then we compute $S_1, S_2, \ldots, S_6$. Having these we employ the homotopy continuation method to determine nontrivial zeros of $J_1(z)$. First we choose an arbitrary vector $A = (A_1, A_2, \ldots, A_6)$, then we follow the solution curve of the homotopy equation (2.2.2) until $t = 1$.

n	Input Vector
1	(.100893900E + 00, -.158603000E - 01)
2	(.334024966E + 03, -.230493231E + 00)
3	(.447584441E + 01, .458962664E + 01)
4	(.266768698E + 05, -.138646669E + 03)
5	(.113324558E + 03, .900923546E + 03)
6	(.245160906E + 07, -.232377074E + 05)

Table (3.1)

This table shows the input Cauchy vector $s = (s_1, \ldots, s_6)$, where s_i is the sum of the ith power of zeros of $J_1(z)$. Table (3.2) is an arbitrary vector in C^6. Finally, the computed zeros of $J_1(z)$ is given in Table (3.3).

n	Arbitrary Chosen Vector A
1	(.580113649E + 00, .950512735E + 00)
2	(.786371143E + 00, .297620264E + 00)
3	(.453366999E + 00, .626194161E - 02)
4	(.275736426E + 00, .305650943E + 00)
5	(.689100711E + 00, .382662239E + 00)
6	(.132902705E + 00, .831857903E + 00)

Table (3.2)

Computed Zeros of $J_1(z)$
z(0) = 0. (trivial solution)
z(1) = (.367073133E + 01, .144951670E - 01)
z(2) = (.101561514E + 02, -.104329900E - 01)
z(3) = (-.707456470E + 01, .536099550E - 02)
z(4) = (-.101568583E + 02, .257022866E - 01)
z(5) = (.707868536E + 01, .450648139E - 02)
z(6) = (-.366324131E + 01, -.554922404E - 01)

Table (3.3)

CHAPTER FOUR

EIGENVALUE PROBLEM

§(4.1) Introduction

In this chapter we shall describe an algorithm for finding the eigenvalues of a given $n \times n$ matrix. The homotopy continuation method of chapter two is implemented to approximate the eigenvalues of a matrix. We shall present an efficient recursive formula to find trace (C^k), $(k = 1,2,\ldots,n)$, where C is an $n \times n$ companion matrix.

It has long been known that the eigenvalue of a matrix can be found by solving its characteristic equation. Namely the eigenvalue problem reduces to solving a polynomial, $P(\lambda)$ of degree n with

$$P(\lambda) = \lambda^n + P_1\lambda^{n-1} + \cdots + P_{n-1}\lambda + P_n \qquad (4.1.1)$$

A large number of methods are available for the determination of the zeros of $P(\lambda)$. The most well known and simple classes of methods are iterative methods. For example the Newton's iteration method

$$\lambda_{n+1} = \lambda_n - \frac{P(\lambda_n)}{P'(\lambda_n)} \qquad (4.1.2)$$

is one of the most popular methods. Such methods have a number of difficulties, among them i) the starting point must be fairly close to desired solution, ii) even if we choose a "good" starting point, the rate of convergence may not be as fast as we wish, iii) some polynomials of high degree are highly sensitive to their coefficients.

§(4.2) Method and Theorems

Let us start this section with the following theorem which will be needed later.

Theorem (4.2.1)

Let $A = (a_{ij})$ be an $n \times n$ matrix with $a_{kk-1} \neq 0$, for some $1 < k \leq n$. Then A is similar to a matrix $B = (b_{ij})$ with $b_{kk-1} = 1$ and $b_{kj} = 0$ for $j \neq k-1$.

Proof:

Let us define a matrix $U = (U_{ij})$ as follows:

$$U_{ij} = \begin{cases} \dfrac{1}{a_{kj}} & \text{if } i = k-1,\ j = k-1, \\[2ex] \dfrac{-a_{kj}}{a_{kk-1}} & \text{if } i = k-1,\ j \neq k-1, \\[2ex] \delta_{ij} & \text{if } i \neq k-1,\ j = 1,2,\dots,n\ . \end{cases} \tag{4.2.1}$$

Define $C = AU$, with $C = (C_{ij})$. Then if $j \neq k-1$,

$$c_{kj} = \sum_{i=1}^{n} a_{ki}U_{ij} = a_{kk-1}U_{k-1j} + a_{kj}U_{jj}$$

$$= a_{kk-1}\left(\frac{-a_{kj}}{a_{kk-1}}\right) + a_{kj} = 0 , \qquad (4.2.2)$$

and if $j = k-1$

$$c_{kj} = c_{kk-1} = \sum_{i=1}^{n} a_{ki}U_{ij} = a_{kk-1}U_{kk-1} = 1 . \qquad (4.2.3)$$

The construction of U implies U is invertible, and $U^{-1} = (v_{ij})$ with

$$v_{ij} = \begin{cases} a_{kj} & \text{if} \quad i = k-1,\ j = 1,2,\dots,n, \\ \delta_{ij} & \text{if} \quad i \neq k-1,\ j = 1,2,\dots,n . \end{cases}$$

It is clear that the k<u>th</u> row of AU is the same as the k<u>th</u> row of $U^{-1}AU$. Hence if we define $B = U^{-1}AU$, then A and B are similar and B satisfies the stated condition. This completes the proof.

Let $A = (a_{ij})$ and $a_{nn-1} \neq 0$, then A is similar to a matrix B_1 with the n<u>th</u> row of the form $(0,0,\dots,1,0)$, that is, there exists an invertible matrix U_1 such that

$$B_1 = U_1^{-1}AU_1 = \begin{pmatrix} b^1_{11} & b^1_{12} & \cdots & b^1_{1n} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ b^1_{n-11} & b^1_{n-12} & \cdots & b^1_{n-1n} \\ 0 & 0 & \cdots & 1 \quad 0 \end{pmatrix} . \qquad (4.2.4)$$

Lets assume $b^1_{n-1n-2} \neq 0$, then there exists an invertible matrix U_2 such that

$$B_2 = U_2^{-1}B_1U_2 = U_2^{-1}U_1^{-1}AU_1U_2 = (U_1U_2)^{-1}A(U_1U_2)$$

$$= \begin{pmatrix} b^2_{11} & b^2_{12} & \cdots & & & b^2_{1n} \\ \vdots & & & & & \\ b^2_{n-21} & b^2_{n-22} & \cdots & & & b^2_{n-2n} \\ 0 & 0 & \cdots & 1 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 0 \end{pmatrix} .$$

Suppose the above procedure can be done successfully $n-1$ times (the case which this procedure fails will be discussed later). Then there exists a sequence of matrices $U_1, U_2, \ldots, U_{n-1}$ such that

$$B_{n-1} = (U_1U_2 \cdots U_{n-1})^{-1}A(U_1U_2 \ldots U_n) = U^{-1}AU$$

$$= \begin{pmatrix} b^{n-1}_{11} & b^{n-1}_{12} & \cdots & b^{n-1}_{1n} \\ 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ \vdots & & & \\ 0 & 0 \cdots & 1 & 0 \end{pmatrix} \qquad (4.2.5)$$

with $U = U_1U_2 \ldots U_{n-1}$.

Remark 1.

Suppose after $n-k$ steps of the above procedure we end up with a matrix $B = (b_{ij})$ where $b_{kk-1} = 0$, but still there exists some $j (1 \leq j < k-1)$ with $b_{kj} \neq 0$. In this case we postmultiply and premultiply this matrix by the matrix E (elementary row operation), where

$$E = \begin{pmatrix} 1 &&&&& \\ & 0 &&& 1 & \\ && 1 &&& \\ &&& \ddots && \\ &&&& 1 & \\ & 1 &&& 0 & \\ &&&&& 1 \end{pmatrix} \begin{array}{l} \\ \to j\underline{th} \text{ row} \\ \\ \\ \\ \to k\text{-}1\underline{th} \text{ row} \\ \\ \end{array} \qquad (4.2.6)$$

The effect of this multiplication will not change the similarity, but it will interchange the $j\underline{th}$ and $k-1\underline{th}$ columns, and simultaneously changes the rows with the same numbers. So we can proceed with the above algorithm.

Remark 2.

If after $n-k$ steps of the above procedure we end up with a matrix of the form B with

$$B = \begin{pmatrix} \alpha_{11} & \alpha_{12} & & \cdots & \alpha_{1n} \\ \vdots & & & & \\ 0 & 0 \cdots 0 & & 0 \cdots 0 & \\ & & 1 & & \\ & & & 1 & 0 \end{pmatrix}, \rightarrow k\underline{th} \text{ row}$$

Then B can be viewed as the following

$$B = \begin{pmatrix} A_1 & A_2 \\ 0 & A_3 \end{pmatrix}$$

with

$$A_1 = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1k-1} \\ \vdots & & \\ \alpha_{k-11} & \cdots & \alpha_{k-1k-1} \end{pmatrix}, \quad A_3 = \begin{pmatrix} \alpha_{kk} & \cdots & \alpha_{kn} \\ & I & \vdots \\ & & 0 \end{pmatrix}.$$

Clearly λ is an eigenvalue of A if and only if λ is an eigenvalue of A_1 or A_3. A_3 is already in companion form, and we continue to find the corresponding companion matrix for A_1.

Overall a certain matrix A can be transformed to a companion matrix C with the same set of eigenvalues. In order to find the eigenvalues of C we first state the following lemma [38].

Lemma (4.2.2)

Let A be an $n \times n$ matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, then

$$\operatorname{trace}(A^k) = \sum_{i=1}^{n} \lambda_i^k \qquad k = 1,2,\ldots$$

In the following we will present a recurrence formula to compute the entries of matrix C^k from matrix C^{k-1} for a companion matrix C (here C^i is the ith power of C).

Proposition (4.2.3)

Let

$$C = \begin{pmatrix} c_1 & c_2 & \cdots & c_{n-1} & c_n \\ & & & & 0 \\ & & I & & \vdots \\ & & & & 0 \end{pmatrix}$$

be a companion matrix. Suppose $C^k = (c_{ij}^k)$ (the kth power of C) has been calculated, then the entries of $C^{k+1} = (c_{ij}^{k+1})$ are given by

$$c_{1j}^{k+1} = c_j c_{11}^k + c_{1j+1}^k , \qquad j = 1,2,\ldots,n-1 \qquad (4.2.7)$$

$$c_{1n}^{k+1} = c_n c_{11}^k , \qquad (4.2.7.1)$$

$$c_{ij}^{k+1} = c_{i-1j}^k , \qquad i = 2,3,\ldots,n \quad j = 1,2,\ldots,n . \qquad (4.2.7.2)$$

Proof:

Since $C^{k+1} = C \cdot C^k = C \cdot (C^k_{ij})$, it is clear that

$$C^{k+1}_{ij} = C^k_{i-1j} \qquad i = 2,3,\ldots,n$$
$$j = 1,2,\ldots,n$$

Also since $C^{k+1} = C^k \cdot C$, for $j = 1,2,\ldots,n-1$ we get

$$C^{k+1}_{ij} = \sum_{\ell=1}^{n} C^k_{i\ell}C_{\ell j} = c_j C^k_{11} + C^k_{1j+1} ,$$

and similarly

$$C^{k+1}_{1n} = c_n C^k_{11} .$$

Remark 1.

This proposition enables us to find the entries of C^{k+1} from C^k. Namely we only need to compute the first row of C^{k+1}. Because the second up to nth rows of C^{k+1} are respectively the first up to $n-1$th rows of C^k by (4.2.11.2). Hence C^{k+1} is obtained from C^k at the cost of n multiplication. Therefore $n(n-1)$ multiplication is necessary to find all entries of $C^2, C^3, \ldots, C^n$.

Having found the entries of C^k $(k = 1,2,\ldots,n)$, we compute

$$S_k = \text{trace}(C^k) , \quad k = 1,2,\ldots,n .$$

In order to determine the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of matrix C (and hence the eigenvalues of A) we may solve the following system of equations:

$$\begin{cases} \lambda_1 + \lambda_2 + \cdots + \lambda_n = S_1 \\ \lambda_1^2 + \lambda_2^2 + \cdots + \lambda_n^2 = S_2 \\ \vdots \\ \lambda_1^n + \lambda_2^n + \cdots + \lambda_n^n = S_n \end{cases} \tag{4.2.8}$$

From here on we employ the homotopy continuation method which is described in chapter two of this dissertation.

§(4.3) Numerical results

In this section we shall consider two examples.

Example 1.

Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ -1 & -2 & 1 \\ 1 & 0 & -1 \end{pmatrix} .$$

Since $a_{32} = 0$, we make use of the elementary matrix E where

$$E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} ,$$

and transfer A to a similar matrix B, with

$$B = EAE = \begin{pmatrix} 1 & 3 & 2 \\ 1 & -1 & 0 \\ -1 & 1 & -2 \end{pmatrix}.$$

Let

$$U_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix},$$

then

$$U_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & -2 \\ 0 & 0 & 1 \end{pmatrix}.$$

Hence

$$U_1^{-1} B U_1 = \begin{pmatrix} 4 & 3 & 8 \\ -4 & -6 & -10 \\ 0 & 1 & 0 \end{pmatrix}.$$

Define

$$U_2 = \begin{pmatrix} -\frac{1}{4} & -\frac{3}{2} & -\frac{5}{2} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Then the corresponding matrices C, C^2 and C^3 are as follows.

$$C = (EU_1U_2)^{-1}A(EU_1U_2) = \begin{pmatrix} -2 & 2 & 8 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

$$C^2 = \begin{pmatrix} 6 & 4 & -16 \\ -2 & 2 & 8 \\ 1 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad C^3 = \begin{pmatrix} -8 & -4 & 16 \\ 6 & 4 & -16 \\ -2 & 2 & 8 \end{pmatrix}.$$

Hence we have the following system

$$\begin{cases} \lambda_1 + \lambda_2 + \lambda_3 = -2 \\ \lambda_1^2 + \lambda_2^2 + \lambda_3^2 = 8 \\ \lambda_1^3 + \lambda_2^3 + \lambda_3^3 = 4 \end{cases}$$

At this stage we use the algorithm of chapter two to find λ_1, λ_2, and λ_3. The computed results are shown in table (4.1).

input vectors	S(1) S(2) S(3)	(−0.2E + 01, 0.0E + 0.0) (+0.8E + 01, 0.0E + 0.0) (+0.4E + 01 0.0E + 0.0)
arbitrary choosen eigenvalues	δ(1) δ(2) δ(3)	(.580113649E + 00, .950512735E + 00) (.786371425E + 00, .297620264E + 00) (.453699900E + 00, .626194161E + 00)
constant vector Pa	Pa(1) Pa(2) Pa(3)	(−.382018497E + 01, −.125439494E + 01) (0.783133603E + 01 −.157657305E + 01) (0.500647786E + 01, −.630500849E + 00)
eigenvalues of the given matrix A	λ(1) λ(2) λ(3)	(−.187555035E + 01, .102511753E + 01) (0.175110070E + 01, .238749298E − 11) (−.187555035E + 01, −.102511753E + 01)
output vector S′	S′(1) S′(2) S′(3)	(−.200000000E + 01, −.229505304E − 11) (0.799999999E + 01, 0.873967565E − 11) (0.399999999E + 01, 0.840714165E − 10)

Table (4.1)

Here S_1, S_2, and S_3 are respectively, trace(A), trace(A^2), and trace(A^3). Likewise Pa(1), Pa(2), Pa(3) are trace(M), trace(M^2), and trace (M^3) for an arbitrary choosen matrix M with known eigenvalues δ(1), δ(2), and δ(3). Finally λ(1), λ(2) and λ(3) are the computed eigenvalues of the given matrix A.

Example 2.

In order to illustrate the accuracy of our algorithm, we will find the eigenvalues of a diagonal matrix D,

$$D = \begin{pmatrix} d_1 & & & & \\ & d_2 & & & 0 \\ & & \ddots & & \\ & 0 & & \ddots & \\ & & & & d_{10} \end{pmatrix}$$

where d_i (for $i = 1,2,\ldots,10$) are given in Table (4.2). The corresponding companion matrix C is computed as

$$C = \begin{pmatrix} c_1 & c_2 & \cdots & c_{10} \\ & & & 0 \\ & I & & \vdots \\ & & & 0 \end{pmatrix}$$

where c_i (for $i = 1,2,\ldots,10$) are also given in Table (4.2).

i	1	2	3	4	5	6	7	8	9	10
d_i	1.0	4.0	-1.0	-3.0	5.0	0.0	2.0	-4.0	2.0	3.0
c_i	5.0	30.0	-150.	-273.	1365.	820.	-4100.	-576.	2880.	0.0

Table (4.2)

Since $c_{10} = 0$, the last column of C is identically zero. This implies that zero is an eigenvalue of C. Hence we may delete the last column and last row of C and denote the remaining part by C_1. To compute $\mathrm{trace}(C_1^i)$ (for $i = 1,2,\ldots,9$) we observe the recursion formulas (4.2.7), (4.2.7.1), and (4.2.7.2) and compute all elements below or on the main diagonal of C_1^9. These are shown in Table (4.3). In fact the first 9 rows from the top of this table are the elements of C_1^9, and from the second row up to the tenth rows are the elements of C_1^8, and so on. Overall we form a system of nine equations with unknowns $\lambda_1, \lambda_2, \ldots, \lambda_9$. Then we use the homotopy continuation method of chapter two to solve this system. The computed results are shown in Tables (4.3) to (4.5).

46587905								
8666515	325530							
1733303	0	-48743760						
307450	196053	-9223500	-2626260					
61490	0	-1648647	0	14160510				
10010	11440	-300300	-147147	2732730	496860			
2002	0	-48620	0	399399	0	-1144780		
257	627	-8250	-7370	75075	24024	-225500	-17280	
55	0	-1023	0	7649	0	-21076	0	14400
5	30	-150	-273	1365	820	-4100	-576	2880
1	0	0	0	0	0	0	0	0
0	1	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0
0	0	0	1	0	0	0	0	0
0	0	0	0	1	0	0	0	0
0	0	0	0	0	1	0	0	0
0	0	0	0	0	0	1	0	0
0	0	0	0	0	0	0	1	0

Table (4.3)

Let us choose a matrix M with eigenvalues $\delta(i)$ $(i = 1,2,\ldots,n)$ where $\delta(i)$ is the $i\underline{th}$ row of Table (2.1). In the following table we show the traces of matrices A^i and M^i $(i = 1,2,\ldots,n)$.

i	trace (A^i)	trace (M^i)
1	.50000000E + 01	(.159697E + 00, -.517790E + 01)
2	.85000000E + 02	(.856692E + 02, -.447630E + 01)
3	.12500000E + 03	(.127149E + 03, -.134409E + 01)
4	.13330000E + 04	(.133421E + 04, -.130606E + 00)
5	.31250000E + 04	(.312621E + 04, -.309642E + 00)
6	.25405000E + 05	(.254070E + 05, .695913E + 00)
7	.78125000E + 05	(.781262E + 05, .242440E + 01)
8	.53533300E + 06	(.535332E + 06, .283408E + 01)
9	.19531250E + 07	(.195312E + 07, .233227E + 01)
10	.11982925E + 08	(.119829E + 08, .214617E + 01)

Table (4.4)

The computer results for eigenvalues of matrix A and the corresponding errors are shown in Table (4.5).

i	Eigenvalues of matrix A	corresponding error
1	(.999999E + 00, -.152670E - 10)	1.144069E - 09
2	(.400000E + 01, -.177490E - 09)	2.392691E - 10
3	(-.100000E + 01, -.202179E - 09)	9.909970E - 10
4	(-.300000E + 01, -.279060E - 09)	8.066281E - 10
5	(.499999E + 01, .139802E - 10)	1.433335E - 11
6	(.116469E + 08, .510748E - 10)	1.165636E - 09
7	(.200000E + 01, -.184396E - 09)	2.051873E - 10
8	(-.399999E + 01, .144042E - 10)	1.407780E - 11
9	(-.199999E + 01, .334406E - 09)	6.867270E - 10
10	(.299999E + 01, .320791E - 09)	5.940594E - 10

Table (4.5)

CHAPTER FIVE

EIGENPAIR PROBLEM

§(5.1) Introduction

The eigenpair (eigenvector, eigenvalue) problem for a square matrix $A \in \mathbb{R}^n \times \mathbb{R}^n$ is that of determining a scalar λ and a vector x such that

$$Ax = \lambda x, \qquad x \neq 0 . \tag{5.1.1}$$

The problem is clearly nonlinear since both λ and x are unknown . Since the eigenvalues are the n roots of the characteristic equation

$$\det(A - \lambda I) \equiv P(\lambda) = 0 , \tag{5.1.2}$$

They can be found without reference to any of the eigenvectors. For a given eigenvalue λ, the corresponding eigenvector is a nontrivial solution of the linear system $Ax = \lambda x$.

This chapter is concerned with homotopy continuation method for calculating the complete set of eigenpairs of a symmetric matrix, and we will avoid of finding explicitly the coefficients of $P(\lambda)$ in order to determine the eigenvalues. Instead a special homotopy is introduced and we shall prove that there are exactly n distinct smooth curves which connect trivial solutions to the desired eigenpairs. In fact these curves are solutions of a certain ordinary differential equations with different initial values, and hence they can be followed numerically by any ordinary differential equations solver.

We emphasize the practical importance of not finding explicitly the coefficients of $P(\lambda)$ in order to evaluate the polynomial. All experienced practitioners are aware of the large error that may result from the use of the approximate coefficients of $P(\lambda)$ for calculation of the zeros of the characteristic polynomials.

§(5.2) The Algorithm

We restrict our discussion to symmetric eigenpair problem

$$Ax = \lambda x, \qquad x \neq 0 \ . \tag{5.2.1}$$

Although through a standard tridiagonalization, we may assume, without loss of generality, that the matrix A is a Jacobian matrix with nonzero off-diagonal elements. The eigenpair problem can be thought of as solving a nonlinear algebraic equation

$$f(x,\lambda) = 0 \tag{5.2.2}$$

where

$$f : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}^n$$

is defined by

$$f(x,\lambda) = Ax - \lambda x \ .$$

There are many well-developed methods which can be used to find the non-trivial solutions of $f(x,\lambda) = 0$. For example, suppose that the spectrum $\sigma(A)$ is simple. Then the classical Newton's method and its many improved modification are particularly well suited for solving (5.2.2). Since its higher Frechét derivatives can easily be determined, the second derivative is constant and higher derivatives vanish. For a detailed discussion of this approach, see [7]. Unfortunately there are some disadvantages in using the Newton's method. Among them Newton's method can converge (if it ever converges) to only one eigenpair at a time. That is in order to compute all n eigenpairs of A, we have to restart the iteration by making n suitable guesses. One possible approach solving this problem by homotopy continuation method is to view (5.2.2) as a system of $n+1$ quadratic polynomials in $n+1$ unknowns. Then the special homotopy defined by Chow, Mallet-Paret and York [5] is applicable for solving (5.2.2). However, there are at least $2^{n+1} - n$ curves diverging to infinity which causes a great inefficiency (particulary for large n). Another approach presented by M. Chu to solve this problem [6]. He defined a homotopy:

$$H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$$

by

$$H(x,\lambda,t) = ([D + t(A - D) - \lambda I]x,\ \tfrac{1}{2}(x^t x - 1)), \qquad (5.2.3)$$

where D is an arbitrary diagonal matrix with distinct elements. Applying this homotopy has the following disadvantages; a) if we follow the n distinct curves suggested by M. Chu we may not get all n eigenpairs, since two of these curves may link into a pair of eigenpairs of the form (x,λ) and $(-x,\lambda)$, (so they actually represent one eigenpair). b) to get all eigenpairs we actually must follow $2n$ distinct curves rather than n curves.

In order to remedy this problem and solve (5.2.2) at a reasonable cost, a special homotopy is constructed as follows:

Let D be an arbitrary diagonal matrix with distinct elements on its diagonal. Construct the homotopy equation

$$H : \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^n \times \mathbb{R}$$

defined by

$$H(x,\lambda,t) = ([(1-t)D + tA - \lambda I]x,\ \varepsilon \sum_{i=1}^{n} x_i - t(x^t x)^{\frac{1}{2}} - 1), \qquad (5.2.4)$$

where ε is a small positive number. It is clear that vectors

$$\begin{bmatrix} x(0) \\ \\ \lambda(0) \end{bmatrix} = \begin{bmatrix} e_i/\varepsilon \\ \\ d_i \end{bmatrix} \qquad i = 1,2,\ldots,n$$

are the eigenpairs of $H(x,\lambda,0)$, where e_i is the standard $i\underline{th}$ unit vector and d_i is the $i\underline{th}$ element of the diagonal matrix D. We should mention here that the crucial step in applying the homotopy continuation method is the construction of an appropriate homotopy, such as (5.2.4) so that i) the existence of a curve connecting the trivial solution and desired solution is assured and ii) the numerical work in following this curve has a reasonable cost.

In the next section we shall show that the homotopy equation (5.2.3) guarantees the existance of n distinct smooth curves. Each of them leads from an obvious starting point to a desired eigenpair. Furthermore if a certain curve links to an eigenpair (x,λ), then there is no other curve that may link to $(-x,\lambda)$. These curves are characterized by an explicit ordinary differential equation with distinct initial values, and hence they can be easily followed by any ordinary differential equations solver . Coupled with the large scale matrix techniques, this method can be used to solve eigenvalue problems for sparse matrices [22].

§(5.3) Theorems

In this section we present some theorems which serve as a theoretical basis for our algorithm.

Theorem (5.3.1)

$0 \in \mathbb{R}^n \times \mathbb{R}$ is a regular value for H. In other words, for each $(\bar{x},\bar{\lambda},\bar{t}) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ with $H(\bar{x},\bar{\lambda},\bar{t}) = 0$. The Jacobian $D_{(\bar{x},\bar{\lambda},\bar{t})}H$ has rank $n+1$.

Proof:

Let $(\bar{x},\bar{\lambda},\bar{t}) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ and $H(\bar{x},\bar{\lambda},\bar{t}) = 0$. Observe that

$$D_{(\bar{x},\bar{\lambda},\bar{t})}H =$$

$$\begin{bmatrix} (1-\bar{t})D+\bar{t}A-\bar{\lambda}I & -\bar{x} & (A-D)\bar{x} \\ (\epsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1,\ldots,\epsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n) & 0 & -(\bar{x}^t\bar{x})^{\frac{1}{2}} \end{bmatrix} \tag{5.3.1}$$

Since $H(\bar{x},\bar{\lambda},\bar{t}) = 0$, we have

$$\epsilon \sum_{i=1}^{n} \bar{x}_i - t(\bar{x}^t\bar{x})^{\frac{1}{2}} = 1 \tag{5.3.2}$$

and

$$((1-\bar{t})D + \bar{t}A - \bar{\lambda}I) \cdot \bar{x} = 0 \tag{5.3.3}$$

We claim that the $(n+1) \times (n+1)$ matrix

$$D_{(\bar{x},\bar{\lambda})}H = \begin{bmatrix} (1-\bar{t})D+\bar{t}A-\bar{\lambda}I & -\bar{x} \\ (\epsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1, \ldots, \epsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n) & 0 \end{bmatrix}$$

is of full rank. Since otherwise there exists a vector $(y,\mu)^t$ with $y \in \mathbb{R}^n$, and $\mu \in \mathbb{R}$ such that

$$D_{(\bar{x},\bar{\lambda})}H \cdot \begin{pmatrix} y \\ \mu \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{5.3.4}$$

Thus

$$((1-\bar{t})D+\bar{t}A-\bar{\lambda}I)\cdot y - \mu x = 0 . \tag{5.3.5}$$

This implies

$$\bar{x}^t \cdot ((1-\bar{t})D+\bar{t}A-\bar{\lambda}I)\cdot y = \mu\bar{x}^t\bar{x} .$$

Since A is symmetric, and $\bar{x}$ is orthogonal to the rowspace of the $(1-\bar{t})D+\bar{t}A-\bar{\lambda}I$, we have

$$\mu\|\bar{x}\|_2^2 = \mu\bar{x}^t\bar{x} = (\bar{x}^t \cdot (1-\bar{t})D + \bar{t}A - \bar{\lambda}I) \cdot y)^t$$

$$= y^t \cdot ((1-\bar{t})D+\bar{t}A-\bar{\lambda}I)\cdot\bar{x} = 0 .$$

This implies $\mu = 0$. Therefore

$$((1-\bar{t})D+\bar{t}A-\bar{\lambda}I)\cdot y = 0 . \tag{5.3.6}$$

Since the matrix $(1-\bar{t})D+\bar{t}A-\bar{\lambda}I$ has a simple spectrum ([10], Lemma 6.1), we conclude that the matrix $B = (1-\bar{t})D+\bar{t}A-\bar{\lambda}I$ has a set of orthogonal eigenvectors, say, $z_1, z_2, \ldots, z_n$, with corresponding eigenvalues

$\delta_1, \delta_2, \ldots, \delta_n$. Let

$$\bar{x} = \alpha_1 z_1 + \alpha_2 z_2 + \cdots + \alpha_n z_n ,$$

$$y = \beta_1 z_1 + \beta_2 z_2 + \cdots + \beta_n z_n .$$

Then by (5.3.3), (5.3.6) and orthogonality we get

$$\alpha_i = \beta_i \delta \qquad i = 1,2,\ldots,n .$$

Therefore $y = \delta\bar{x}$ for some δ. Substituting $y = \delta\bar{x}$ and $\mu = 0$ in (5.3.4) we get

$$\begin{bmatrix} (1-\bar{t})D + \bar{t}A - \bar{\lambda}I & -\bar{x} \\ \varepsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_1, \ldots, \varepsilon - t(\bar{x}^t\bar{x})^{-\frac{1}{2}}\bar{x}_n & 0 \end{bmatrix} \begin{bmatrix} \delta\bar{x} \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} .$$

Thus by (5.3.2) we have,

$$0 = \delta(\varepsilon \sum_{i=1}^{n} \bar{x}_i - t(\bar{x}^t x)^{-\frac{1}{2}}) = \delta . \qquad (5.3.7)$$

Hence $D_{(\bar{x},\bar{\lambda})}H$ is of rank $n+1$. This completes the proof.

Remark.

We have restricted our discussion with the Jacobi structure of the matrix A. This is needed only as a sufficient condition for Theorem (5.3.1). This condition may be rephrased as "choosing D so that the matrix $(1-t)D + tA$ has a simple spectrum for any $t \in [0,1]$". Overall, it is only needed that the matrix $D_{(x,\lambda,t)}H$ be of full rank for any $(x,\lambda,t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}$ with $H(x,\lambda,t) = 0$. Apparently the sparse matrix techniques can be incorporated in any of these cases.

As is the usual procedure of the homotopy continuation method we start from a trivial solution of $H(\cdot,\cdot,0)$ at $t = 0$, and follow the generated path as t increases from zero to one. We hope the trivial eigenpairs deform into the eigenpairs of the original matrix A. Hence we would be able to follow the n distinct connected paths from the trivial system to the original problem. In order to assure that this process works, we prove the following:

Theorem (5.3.2)

Let us define

$$\Gamma = \{(x,\lambda,t) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R} : H(x,\lambda,t) = 0\} .$$

a) Γ is a one dimensional smooth manifold,

b) as t increases, the curve Γ will never turn back.

Proof:

Part (a) is in fact a standard result from the differential topology [35]. That is, a repeat use of the Implicit Function Theorem implies that Γ consists of one dimensional manifold.

In order to prove (b), let Γ be parameterized with a parameter, θ. Along each component, we may take the derivative with respect to the parameter θ. The set Γ is then characterized by

$$D_{(t)}H \cdot \frac{dt}{d\theta} + D_{(x,\lambda)}H \cdot \begin{bmatrix} \frac{dx}{d\theta} \\ \frac{d\lambda}{d\theta} \end{bmatrix} = 0 \ . \qquad (5.3.8)$$

We claim $\frac{dt}{d\theta} \neq 0$, since otherwise

$$D_{(x,\lambda)}H \cdot \begin{bmatrix} \frac{dx}{d\theta} \\ \\ \frac{d\lambda}{d\theta} \end{bmatrix} = 0 \ .$$

Hence $D_{(x,\lambda)}H$ would be singular. This completes the proof.

Now consider starting a path at $t = 0$. Since $\dot{t} \neq 0$. without loss of generality, we may assume $\dot{t} > 0$. So as we increase the parameter θ, $t(\theta)$ cannot reverse. The following lemma shows that as t goes to one the curve Γ remains bounded. In other words for any $0 < t_0 < 1$, the set

$$\Gamma_0 = \{(x,\lambda) \in \mathbb{R}^n \times \mathbb{R} : H(x,\lambda,t) = 0 \text{ for some } t \in [t_0,1]\} \qquad (5.3.9)$$

is bounded.

<u>Lemma (5.3.3)</u>

The set Γ_0 is a bounded set.

Proof:

By equation (5.3.3)

$$|\lambda(t)|\,\|x\| = \|[(1-t)D+tA]x\| \leq \|(1-t)D+tA\|\,\|x\|$$
$$\leq (\|D\|+\|A\|)\,\|x\|$$

Hence

$$|\lambda(t)| \leq \|D\|+\|A\| \quad . \tag{5.3.10}$$

Also for any $(x,\lambda) \in \Gamma_0$, equation (5.3.2) implies

$$t_0\|x\| \leq \|x\| = |1+\varepsilon \sum_{i=1}^{n} x_i|$$
$$\leq 1+\varepsilon \sum_{i=1}^{n} |x_i| = 1+\varepsilon\|x\|_1$$
$$\leq 1+\varepsilon\sqrt{n}\|x\| \quad .$$

Thus for ε small

$$\|x\| \leq \frac{1}{t_0-\varepsilon\sqrt{n}} \quad . \tag{5.3.11}$$

Therefore Γ_0 is bounded. (Here by $\|\cdot\|$ we mean $\|\cdot\|_2$.)

Part (b) of Theorem (5.3.2) implies that Γ will never turn back. Thus Γ can be parametrized by the variable t. Then (5.3.8) becomes

$$\begin{bmatrix} \frac{dx}{dt} \\ \\ \frac{d\lambda}{dt} \end{bmatrix} = \begin{bmatrix} (1-t)D+tA-\lambda I & -x \\ \\ \varepsilon-t(x^tx)^{-\frac{1}{2}}x_1,\ldots,\varepsilon-t(x^tx)^{-\frac{1}{2}}x_n & 0 \end{bmatrix}^{-1} \begin{bmatrix} (D-A)x \\ \\ -(x^tx)^{\frac{1}{2}} \end{bmatrix} \tag{5.3.12}$$

$$\begin{bmatrix} x(0) \\ \\ \lambda(0) \end{bmatrix} = \begin{bmatrix} e_i/\varepsilon \\ \\ d_i \end{bmatrix} \qquad i = 1,2,\ldots,n \; .$$

These differential equations may be solved by any ordinary differential equations solver. Over all for each $1 \leq i \leq n$ the numerical solution of (5,3,12) at $t = 1$ is an approximation for an eigenpair of the matrix A. The following lemma assures that among the computed eigenpairs we will not have a pair of eigenpairs of the form $[x,\lambda]^t$ and $[-x,\lambda]^t$.

Lemma (5.3.4)

There is no pair of eigenpairs of the form $[x,\lambda]^t$, $[-x,\lambda]^t$.

Proof:

Since otherwise both vectors satisfy equation (5.3.2) with $t = 1$. Subtracting the two resulting equations gives $\sum_{i=1}^{n} x_i = 0$. Substituting this in (5.3.2) yields $\|x\|_2 = -1$, which is not true.

§(5.4) The Algorithm

We have seen that the set

$$\Gamma = \{(x,\lambda,t) : H(x,\lambda,t) = 0\} , \qquad (5.4.1)$$

consists of exactly n curves, and each of them is the solution of the differential equations (5.3.12). We proceed as follows:

I) Choose any diagonal matrix D with distinct elements on its diagonal.

II) Set $i = 0$.

BIBLIOGRAPHY

BIBLIOGRAPHY

1. Abraham, R. and Robbin, J. (1967). Transversal mapping and flows, Benjamin, New York.

2. Allgower, E. and Georg, K. (1980). Simplicial and continuation methods for approximating fixed-points and solutions to systems of equations, SIAM Review, vol. 22, No. 1.

3. Bauer, F.L., Rutishauser, H. and Stiefel, E. (1963). New aspects in numerical quadrature, Proc. Sympos. Appl. Math. Vol. 15, pp. 199-218. Amer. Math. Soc. Providence, R.I., MR30 #4384.

4. Chow, S.N., Mallet-paret, J. and Yorke, J.A. (1978). Finding zeros of maps: Homotopy methods that are constructive with probability one, Math. Comput., 32, pp. 887-899.

5. __________, ____________, ______________ (1979). A homotopy method for locating all zeros of a system polynomials, Functional differential equations and approximation of fixed points, Springer lecture notes in mathematics, No. 730, pp. 77-88.

6. Chu, M.T. (1983). A simple application of the homotopy method to symmetric eigenvalue problems, "to appear".

7. Collatz, L. (1966). Functional analysis and numerical mathematics, Acd. Press. New York.

8. Davidenko, D. (1953a). On a new method of numerically integrating a system of nonlinear equations, Dokl. Akad. Nauk SSSR, 88, pp. 601-604. (in Russian).

9. ___________, (1953b). On the approximation solution of a system of nonlinear equations, Ukrain, Math. Ž. 15, pp. 190-206 (in Russian).

10. Deift, P., Lund, F. and Trubowitz, E. (1980). Nonlinear wave equation and constrained Harmonic motion, Comm. Math. Phys. 74, pp. 141-188.

11. Delves, L.M. and Lyness, T.N. (1967). A numerical method for locating the zeros of an analytic function, Math. Comp., vol. 21, pp. 543-560.

12. Drexler, F.J. (1977). Eine method zur Berechnung Sämtlicher l'sunger von polynomgleichunges systemen, Numerische Mathematic 29, pp. 45-58.

13. ____________, (1978). A homotopy method for calculating of zeros of zero-dimensional polynomial ideals, in continuation methods, ed. H. Wacker, New York: Academic Press.

14. Eaves, B.C. (1971a). On the basic theory of complementary, Math. Programming, 1.1, pp. 68-75.

15. __________, (1972). Homotopy for computation of fixed point, Math Programming, Ibid., 3.1, pp. 1-22.

16. __________, (1976). A short course in solving equations with PL homotopies, SIAM-AMS Proc., 9, pp. 73-143.

17. Friedland, S. (1977). Inverse eigenvalue problems, Linear algebra and Appl., 17, pp. 15-51.

18. Garcia, C.B. and Gould, F.J. (1978). A theorem on homotopy paths, Mathematics of operational research, vol. 3, No. 4.

19. ____________, and Li, T.Y. (1980). On the number of solutions of polynomial system of equation, SIAM Journal of numerical analysis, No. 17, pp. 540-546.

20. ___________, and Zangwill, W.I. (1978). Global continuation methods for finding all solutions to polynomial systems of equations in n variables, in: Symposium on external methods and systems analysis, Springer Verlag, Berlin.

21. ____________, ________________. (1979). Determining all solutions to certain systems of nonlinear equations, Mathematics of operations Research, vol. 4, No. 1.

22. ____________, ________________. (1979). An approach to homotopy and degree theory, Mathematics of operations Research, vol. 4, No. 4, pp. 390-405.

23. Hirsch, M. (1963). A proof of the nonretractability of a cell onto its boundary, Proc. Amer. Math. Soc. 14, pp. 364-365.

24. Karamardian ed. (1977). Fixed points algorithms and applications, Academic Press, New York.

25. Kellogg, R.B., Li, T.Y. and Yorke, J.A. (1976). A constructive proof of the Brouwer fixed-point Theorem and computational results, SIAM J. Numer. Anal., vol. 13, No. 4.

26. ____________, ________, ______________. (1977). A method of continuation for calculating a Brouwer fixed point, computing fixed points with applications, Kuramardian, ed., Academic Press, New York.

27. Kuhn, H. and Mackinnon, J.G. (1975). Sandewich method for finding fixed points, J. optimization Theory Appl., 17, pp. 189-204.

28. Lahaye, E. (1934). Une méthode de résolution d'une catégorie d'équations trancendantes, C.R. Acad. Sci., Paris, 198, pp. 1840-1842.

29. _________, (1935). Sur la représentation des racines systemes d'èquations transcendantes, Deuxieme congres national de sciences, 1, pp. 141-146.

30. Lemke, C.E. and Howson, J.T. (1964). Equilibrium points of bimatrix games, Siam J. Applied Math., 12, 2, pp. 413-423.

31. Li, T.Y. (1976). Computing the Brouwer fixed point by following the continuation curve, Proc. Sem. Dalhousia Univ., Halifax, Academic Press, New York, pp. 131-135.

32. _______, (1982). On locating all the zeros of an analytic function within a bounded domain, "to appear".

33. Lyness, J.N. and Delves, L.M. (1967). On numerical contour integration round a close contour, Math. Comp. vol. 21, pp. 561-577.

34. Merrill, O.H. (1972). Applications and extensions of an algorithm that computes fixed points of a certain upper semi-continuous point to set mapping, Ph.D. Thesis, Dept.

35. Milnor, J. (1965). Topology from the differentiable viewpoint, Univ. Virginia Press, VA.

36. Morris, J. Li. (1983). Computational methods in elementary numerical analysis, John Wiley and Sons.

37. Noether, E., Van der Waerden, B.L. (1928). Die alternative bei nichlinearen Gleichvgen, Nachrichten der Gesselschaft der Wissenschaften zu Göttingen, Math. Phys. Klass, pp. 77-87.

38. Ortega, J.M. and Rheinboldt, W.C. (1970). Iterative solution of nonlinear equations in several variables, Academic Press, New York - London.

39. Raleston, A. and Rabinowitz, P. (1978). A first course in numerical analysis, McGraw-Hill, Inc.

40. Saigal, R. (1977b). On the convergence rate of algorithms for solving equations that are based on complementarity pivoting, Math. Oper. Res., 2, pp. 108-124.

41. Scarf, H. (1967). The approximation of fixed points of a continuous mapping, Siam J. Appl. Math., 15,5, pp. 1328-1343.

42. Shampine, L.F. and Gordon, M.K. (1975). Computer solution of ordinary differential equations: the initial value problem, Freeman and company, San Francisco.

43. Van der Waerden, B.L. (1931). Modern algebra, I and II, New York: F. Ungar Publishing.

44. Walker, H., ed. (1978). Continuation methods, Academic Press, New York.

45. Wilkinson, J.H. (1965). The algebraic eigenvalue problem, Clarendon Press, Oxford.