

To the Graduate Council:

I am submitting herewith a thesis written by Nancy G. Saltzman entitled "Alternating Direction Implicit Iteration for Iteration Matrices with Some Complex Spectra." I have examined the final copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Mathematics.

Eugene L. Wachspress
Eugene L. Wachspress, Major Professor

We have read this thesis
and recommend its acceptance:

Lawrence A. Bales

James Kosatka

Accepted for the Council:

Lowminkel
Vice Provost
and Dean of The Graduate School

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a Master's degree at The University of Tennessee, Knoxville, I agree that the Library shall make it available to borrowers under rules of the Library. Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from or reproduction of this thesis may be granted by my major professor, or in his absence, by the Head of Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Signature

Nancy B. Johnson

Date

Mar 5th, 1987

ALTERNATING DIRECTION IMPLICIT ITERATION
FOR ITERATION MATRICES WITH
SOME COMPLEX SPECTRA

A Thesis

Presented for the

Master of Science

Degree

The University of Tennessee, Knoxville

Nancy G. Saltzman

June 1987

To S.P.E., S.S., and D.S.S.

ACKNOWLEDGEMENTS

I wish to thank my advisor, Eugene Wachspress, for suggesting the topic of this thesis, and for providing groundwork, direction, aid, and encouragement for my investigations of it. I wish to thank Allen Crider for having written most of the computer code I needed. Initial effort by Allen Crider and Eugene Wachspress was supported by a grant from the Electromechanics Branch of the General Electric Corporate Research and Development Center, Schenectady, New York.

ABSTRACT

Alternating Direction Implicit iteration was introduced by Peaceman and Rachford to solve parabolic and elliptic linear systems. The theory of optimum iteration parameters for iteration matrices with real spectra was developed using Chebyshev minimax theory. Using this method to solve the Lyapunov matrix equation arising from the computations by Hurwitz of impedance boundary conditions for finite element problems leads to consideration of the case of certain spectra with small complex components. Rouché's theorem provides the basis for this generalization. A theory to find the optimum set of parameters for a complex spectrum among all sets optimum for real spectra is developed. Theoretical error bounds are derived much as for the real spectral case. Numerical tests indicate that this theory is appropriate only for very small complex spectral components. Then actual convergence rates are slightly improved but theoretical error bounds are much more valid. These optimum iteration parameters converge towards one with increasing complex spectral component. Since Wachspress has shown the optimum iteration parameters for circular spectra to be one, repeated, it is expected that further investigation will show optimum iteration parameters converging as spectra vary from strictly real to complex and circular.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
II. LITERATURE REVIEW	4
III. THEORY OF OPTIMUM PARAMETER SELECTION	13
A New View of Chebyshev Acceleration for Some Complex Spectra	13
Lemniscates and Their Approximating Ellipses For Analysis of ADI Iteration	19
An Estimate of the Effective Spectral Radius Of ADI Iteration	29
IV. COMPUTER IMPLEMENTATION AND RESULTS	32
Methods	32
Tests and Results	39
Relation to the Theory for Eigenvalues On a Circle	50
V. CONCLUSION	56
LIST OF REFERENCES	59
APPENDICES	61
APPENDIX A	62
APPENDIX B	63
APPENDIX C	64
VITA	73

LIST OF FIGURES

FIGURE	PAGE
III-1. The 1-lemniscates of $T_t(z')$	15
III-2. The γ -lemniscates of $T_t(z')$	18
III-3. The 1-lemniscates of $\text{sn}[K_2(2t\omega+1), k_2]$	22
III-4. The γ -lemniscates of $\text{sn}[K_2(2t\omega+1), k_2]$	23
III-5. The conservative regions of the lemniscates shown in Figure III-4	25
IV-1. A rectangular spectrum R_s normalized so its real extent is ln-symmetric, shown with the association region $\Lambda_{s_{\text{opt}}} = \text{dn}(\omega K, k)/\sqrt{k'}$	34
IV-2. Natural logarithm of a ln-symmetrized rectangular spectrum R_s with three sample regions E.	37
IV-3. Actual error and theoretical error bounds for spectra in six rectangles.	43
IV-4. Actual error and theoretical error bounds for spectra in the two six-sided regions a and b and the two ten-sided regions a and b in Table 1.	48
IV-5. Actual error and theoretical error bounds after 20 iterations for spectra in the ten- sided region a in Table 1 for varying k'	51
IV-6. Qualitative illustration of expected behavior of optimum iteration parameters.	54

CHAPTER I

INTRODUCTION

The first Alternating Direction Implicit (henceforth ADI) iterative method of solving parabolic and elliptic partial differential equations was introduced in 1955 by Peaceman and Rachford. Since then the method has been modified and improved and some theoretical understanding of it has been developed. The application of this method to solve the Lyapunov matrix equation¹ arising from the computations by Hurwitz [1984] of impedance boundary conditions for finite element problems led to this consideration of the ADI method in the case when the iteration matrix has complex eigenvalues. No investigation of the ADI method with complex eigenvalues has previously been published, although some of the results presented here appear in the Proceedings of the 1986 ACM meeting in Dallas (Ellner and Wachspress [1986]).

In this thesis ADI iteration for spectra with small complex components relative to the real components is considered theoretically and this theory is tested with some numerical examples. Our goal is not to seek optimum parameters for given regions but to analyze domains for which existing sets of parameters (which are optimal for real spectra) are optimal. These regions have positive real part and are approximately such that $1/\lambda$ is in the region

whenever λ is. Theoretical bounds on convergence of the ADI method for these regions are calculated. Approximately optimal parameters are found for more general spectral regions as follows, where the parameters are approximately optimal in the sense that the spectra approximate the regions on which the parameters are optimal. Key eigenvalues of a spectrum define the minimal convex polygon which contains the spectrum. This polygon can be fit into the domain for which optimum parameters are known which gives the best convergence.

It is important to note that henceforth where confusion is likely the term "true optimum parameters" will mean parameters optimal over all conceivable parameter choices. The term "optimum parameters" will mean parameters optimum over those sets of parameters which are true optimum parameters for some real spectrum. When only real spectra are being considered, as in the literature review in Chapter II, optimum parameters are the same as true optimum parameters.

Results of the numerical experiments done to test the theory indicate that it is valid only for spectra with very small complex components. For these spectra, the actual convergence rates obtained are improved by using iteration parameters generated by this theory, but the improvement is fairly small. However, the theoretical error bounds on

convergence are very reliable, while those using the previous theory for real spectra are overly optimistic.

Chapter II contains some history of the ADI method and a review of the relevant literature pertaining to the ADI method and to the Lyapunov matrix equation. It is based primarily on similar reviews in recent papers by Wachspress. The theory of the specific case under investigation is presented in Chapter III. The first section of Chapter III reviews the theoretical background for Chebyshev acceleration from the perspective which will be used in the new analysis. Section 2 applies these ideas to find the complex spectral regions for which parameters optimal for real spectra are truly optimal. In Section 3 the theoretical error bound for spectra in these regions is derived. The first section of Chapter IV explains the method used in implementing the theory for computing, that is, how to fit a convex polygonal spectral region into the region with truly optimum parameters for which the convergence will be fastest. The results of the implementation are described in Section 2 and related to the previously developed theory for circular spectra in Section 3. Chapter V contains a brief conclusion.

CHAPTER II

LITERATURE REVIEW

The most recent publication on the application of ADI iteration was a summary for conference proceedings by Ellner and Wachspress [1986]. That paper contains much background on ADI iteration (taken largely from Wachspress [1984]) and on the Lyapunov matrix problem. This chapter draws heavily on the appropriate sections in that paper, presenting much of the same material in much the same way.

The first ADI scheme was suggested by Peaceman and Rachford in 1955 to solve the heat flow equation $\frac{\partial u^2}{\partial x^2} + \frac{\partial u^2}{\partial y^2} = \frac{\partial u}{\partial t}$ and Laplace's equation $\frac{\partial u^2}{\partial x^2} + \frac{\partial u^2}{\partial y^2} = 0$ with boundary conditions. In this method, the real symmetric-positive-definite (henceforth SPD) coefficient matrix A , which arises from the five-point discretization of the problem into the form

$$Au = b \tag{1.1}$$

is split into two components, the horizontal part H which arises from the discretization of the derivatives with respect to x , and the vertical part V which arises from the discretization of the derivatives with respect to y . The two step iteration scheme which results, called Peaceman-Rachford iteration, is

$$\begin{aligned}
(H + \omega_j I) u_{j-\frac{1}{2}} &= -(V - \omega_j I) u_{j-1} + b \\
(V + \omega_j I) u_j &= -(H - \omega_j I) u_{j-\frac{1}{2}} + b
\end{aligned}
\tag{1.2}$$

The convergence of this method and computation of optimum parameters ω_j for it were investigated theoretically by Wachspress [1962, 1963, 1966, and 1968] for the case where H and V commute, called the model problem. Wachspress established the existence and uniqueness of optimum iteration parameters which are expressed as a function of the eigenvalue spectra of H and V . Optimum parameters can be easily computed when the number of iterations is an integral power of two by a method using a folding technique in some ways similar to the FFT. Jordan (see Wachspress [1963 and 1984]) used elliptic functions to find optimum parameters for any t .

The Peaceman-Rachford iteration was generalized in two significant ways in the 1960's. Wachspress [1966] discovered that the effects of renormalization of H and V to enhance the efficiency of the iteration could be obtained by replacing $\omega_j I$ by $\omega_j F$ in equations (1.2), where F is a diagonal matrix. The condition for having a model problem is then $HF^{-1}V = VF^{-1}H$.

Peaceman, Rachford, and Douglas found that the iteration scheme (1.2) can diverge when a different value for the iteration parameter is used in both sweeps, whereas

convergence is assured when the same value is used in both sweeps. However, Wachspress [1963] was able to find optimum iteration parameters which may vary with each sweep as a function of the distinct spectral intervals of H and V (or of HF^{-1} and VF^{-1} when the parameters are multiplied by F) which can be a great improvement over the previous method, even to the point of allowing H and V to be singular or have negative eigenvalues as long as A is SPD.

The theory of Jordan (see Wachspress [1966]) for finding optimum parameters for any number of iterations t is more closely related to this study. Jordan reduced the problem of optimum parameter selection to a minimax problem solved by Zolotareff using results of Chebyshev in 1877. This theory was applied to electric filter design by Cauer in 1934. In order to minimize the spectral radius of the ADI iteration the ω_j must be chosen to minimize the maximum absolute value of the function

$$Q_t(\lambda, \gamma, \omega) = \prod_{j=1}^t \frac{(\lambda - \omega_{v_j})(\gamma - \omega_{h_j})}{(\lambda + \omega_{h_j})(\gamma + \omega_{v_j})} \quad (1.3)$$

where λ and γ vary over the spectra of H and V , respectively (which are real). Jordan transformed the variables λ and γ so they both have the same range and the problem can be written as a function of one variable x

$$Q_t(x, \omega) = \prod_{j=1}^t \frac{(x - \omega_j)}{(x + \omega_j)} \quad (1.4)$$

$k' \leq x \leq 1$. He then substituted $x = \operatorname{dn}(Ku, k)$, where Ku is

the argument and k the modulus of the elliptic function dn , K being the complete elliptic integral of the first kind of modulus k . The solution is to choose ω_j such that

$$Q_t(x, \omega_j) = \prod_{j=1}^t \frac{\text{dn}(Ku, k) - \omega_j}{\text{dn}(Ku, k) + \omega_j} = \frac{\text{dn}(tK_1 u, k_1) - \sqrt{k'_1}}{\text{dn}(tK_1 u, k_1) + \sqrt{k'_1}}, \quad (1.5)$$

where $\frac{K'_1}{K_1} = \frac{tK'}{K}$.

since this function satisfies the Chebyshev alternance property and is therefore the unique function with smallest maximum absolute value on the interval $k' \leq x \leq 1$. The choice of ω_j is accomplished through two properties of elliptic functions. First, a rational function of elliptic functions is itself an elliptic function. Second, if two elliptic functions have the same periods, poles, and zeroes, then their ratio is a constant. This theory and the results are presented in detail by Wachspress [1966]. The results can also be derived from the elliptic sn function using the identity

$$\sqrt{k_2} \text{sn}[K_2(2u+1), k_2] = \frac{\text{dn}(K_1 u, k_1) - \sqrt{k'_1}}{\text{dn}(K_1 u, k_1) + \sqrt{k'_1}}, \quad \sqrt{k_2} = \frac{1 - \sqrt{k'_1}}{1 + \sqrt{k'_1}}. \quad (1.6)$$

This form turned out to be most useful for this study.

Ellner and Wachspress [1986] give the following description, paraphrased here, of two recent generalizations. An important application of model problem ADI iteration is as a preconditioner for compound iteration. It

can be an improvement over a Laplacian preconditioner for conjugate gradient or Chebyshev iteration for the diffusion equation. This occurs when the diffusion equation has a diffusion coefficient which is a separable function of x and y and therefore the eigenfunctions of the equation are separable. Then the five-point difference equations of the box integration method applied to the problem yield a model problem for ADI iteration. This can be extended to include non-separable functions when they can be approximated by separable ones. Second, a transformation of variables was discovered that gives an ADI model problem for nine-point equations. These nine-point equations arise in solving finite element problems over rectangular grids.

The subject of the ADI iteration application in this thesis is a specific type of Lyapunov matrix equation. The general Lyapunov matrix problem is to find an $m \times n$ matrix X such that

$$AX + XB = C \quad (1.7)$$

where A is an $m \times m$ matrix, B an $n \times n$ matrix, and C an $m \times n$ matrix. A unique solution exists when $s+q \neq 0$ for all pairs s, q of eigenvalues of A and B , respectively.

Rothschild and Jameson [1970] described four important methods developed in the 1960's for solving the Lyapunov matrix equation

$$A^T X + XA = C, \quad (1.8)$$

when all matrices are $n \times n$. These are the methods of Geiss, et al., Ma, Smith, and Jameson, with numbers of arithmetic operations, respectively, $O(n^6)$, $O(n^5)$, $O(n^5)$, and $O(n^4)$. Bartels and Stewart [1972] developed an algorithm for solving equation (1.8) with $O(N^4)$ arithmetic operations where N is the larger of n , m . This method uses the QR algorithm to compute the real Schur decompositions of the matrices A and B to upper quasi-triangular matrices (that is, triangular with possible non-zero 2×2 blocks along the diagonal). Golub, Nash, and Van Loan [1979] improved on this method by substituting the decomposition of A to upper Hessenberg form by Householder's method.

The Lyapunov equation is suitable for solution by ADI iteration when A and B are N -stable, that is, the real parts of their eigenvalues are positive, and, more generally, when the real part of $s+q$ is positive for all pairs s, q of eigenvalues of A and B , respectively (see Ellner and Wachspress [1986]). The method published in 1968 by Smith applies when A and B are N -stable and is the basis of the iterative method used in this thesis. The following concise summary of Smith's method and description of its generalization is due to Wachspress (see Ellner and Wachspress [1986]).

The identity

$$(qE+A)X(qI+B) - (qE-A)X(qI-B) = 2qC \quad (1.9)$$

holds for E and I the identity matrices of dimension $m \times m$ and $n \times n$, respectively, and q any positive scalar. Multiplying this equation on the left by $(qE+A)^{-1}$ and on the right by $(qI+B)^{-1}$ gives

$$X - UXV = W \quad (1.10)$$

where

$$\begin{aligned} U &= (qE+A)^{-1}(qE-A) \\ V &= (qI-B)(qI+B)^{-1} \\ W &= 2q(qE+A)^{-1}C(qI+B)^{-1}. \end{aligned} \quad (1.11)$$

A formal solution to equation (1.10) is

$$X = \sum_{k=1}^{\infty} U^{k-1} W V^{k-1}. \quad (1.12)$$

This series is convergent since the spectral radii of U and V are less than unity. The sequence of matrices Y_r defined by

$$Y_{r+1} = U^{2^r} Y_r V^{2^r} + Y_r, \quad Y_0 = W \quad (1.13)$$

approaches X rapidly, and U^{2^r} and V^{2^r} are obtained by squaring $U^{2^{r-1}}$ and $V^{2^{r-1}}$, respectively. This iteration is quadratically convergent with each step requiring four matrix multiplications and one matrix addition, and computation arithmetic $O(N^3)$.

Although the generalization of Smith's iterative solution still has $O(N^3)$ arithmetic operations, fewer iterations are required and A and B need not be N -stable, although the real part of $p+q$ must be positive for all

pairs p, q of eigenvalues of A and B , respectively. Also, a parallelism is introduced for efficient use of an array computer. Equation (1.9) is generalized to

$$(pE+A)X(qI+B) - (qE-A)X(pI-B) = (p+q)C, \quad (1.14)$$

then multiplied on the left by $(pE+A)^{-1}$ and on the right by $(qI+B)^{-1}$ to get

$$X - U(p, q)XV(p, q) = W(p, q) \quad (1.15)$$

where

$$\begin{aligned} U(p, q) &= (pE+A)^{-1}(A-qE) \\ V(p, q) &= (B-pI)(qI+B)^{-1} \\ W(p, q) &= (p+q)(pE+A)^{-1}C(qI+B)^{-1}. \end{aligned} \quad (1.16)$$

Smith's iteration, equation (1.13), becomes

$$\begin{aligned} Y_{r+1}(p, q) &= U^{2^r}(p, q) Y_r(p, q) V^{2^r}(p, q) + Y_r(p, q), \quad (1.17) \\ Y_0(p, q) &= W(p, q) \end{aligned}$$

where Y_r is related to the solution X by

$$Y_r(p, q) = X - U^{2^r}(p, q)XV^{2^r}(p, q). \quad (1.18)$$

The p and q should be chosen so that the spectral radii of $U(p, q)$ and $V(p, q)$ are as small as possible, but less than unity for convergence. This is the ADI minimax problem for $t=1$.

Prof. G.J. Habetler (R.P.I.) noted that the model problem for ADI iteration, equation (1.1) where V and H commute, is related to the Lyapunov equation $AX+XB = C$ by defining $H(X) = AX$ and $V(X) = XB$. It is not necessary that A and B commute for model problem theory to apply. Although

model problem theory was developed for H and V with real spectra, this condition also is not necessary.

The particular problem which motivated this thesis comes from the "infinitesimal scaling" method of H. Hurwitz [1984] for computing impedance boundary conditions for finite element problems. Again following the summary of Ellner and Wachspress [1986], given real SPD matrices M and N and real antisymmetric matrix Q , then SPD matrix C is found as an impedance boundary condition from the non-linear matrix equation

$$(C+Q)N^{-1}(C+Q^T) = M. \quad (1.19)$$

If $A_n = (C_n+Q)N^{-1}$, a quadratically convergent Newton-type iteration for C is

$$A_n C_{n+1} + C_{n+1} A_n = M - QN^{-1}Q^T + C_n N^{-1}C_n. \quad (1.20)$$

This is the Lyapunov equation (1.8) with A N -stable since its symmetric part is SPD. This problem is particularly well suited for ADI iteration because the norm of Q is very small compared to the norm of C so the spectrum lies in a rectangle in the complex plane with strictly positive real part and comparably small complex part symmetric about the real axis.

CHAPTER III

THEORY OF OPTIMUM PARAMETER SELECTION

1. A New View of Chebyshev Acceleration

For Some Complex Spectra

In this section we consider the existing theory of Chebyshev acceleration for some complex spectra from a new perspective. We will call a function $F(z)$ in a class of functions $\mathcal{F}$ an optimal function of that class on a region S in the complex plane $\mathbb{C}$ if it satisfies

$$\max_{z \in S} |F(z)| = \min_{G \in \mathcal{F}} \max_{z \in S} |G(z)| \quad (1.1)$$

where $|\cdot|$ denotes the absolute value. Rather than the usual approach of looking for optimal functions on a given region containing the eigenvalues, we investigate the regions on which a certain class of functions is optimal. This approach proves fruitful when applied to the ADI problem with spectra with small complex components. We will make use of the lemniscate of a function which is continuous on a domain, similar to the lemniscate defined by Opfer and Schober [1984].

Definition: The γ -lemniscate of a function $F(z)$ continuous on a region $\Gamma \subset \mathbb{C}$ is

$$L = \{z \in \Gamma: |F(z)| \leq \gamma\}. \quad (1.2)$$

The theory is well developed in the case of real eigenvalues. For eigenvalues in the x -interval $a \leq x \leq b < 1$, the optimal t -degree polynomial is the scaled Chebyshev polynomial

$$P_t(x') = T_t(x')/T_t(\mu) \quad (1.3)$$

where $x' = [2x - (a+b)]/(b-a)$, $-1 \leq x' \leq 1$, and $\mu = [2 - (a+b)]/(b-a) > 1$. The error reduction for t Chebyshev iterations is

$$\max_{-1 \leq x' \leq 1} |P_t(x')| = 1/T_t(\mu). \quad (1.4)$$

The real line segment is only part of the region for which this is the optimal polynomial with $\max_{-1 \leq x' \leq 1} |P_t(x')| = 1/T_t(\mu)$. When complex eigenvalues with the same real range are allowed, the same polynomial $P_t(z') = T_t(z')/T_t(\mu)$, now as a function of a complex variable, is the optimal polynomial on its $1/T_t(\mu)$ -lemniscate. This lemniscate in the z' -plane, shown in Figure III-1b, oscillates between the degenerate ellipse which is the line segment $-1 \leq z' \leq 1$ and the ellipse

$$\left[\operatorname{Re} \left[\frac{z'}{\cosh v} \right] \right]^2 + \left[\operatorname{Im} \left[\frac{z'}{\sinh v} \right] \right]^2 = 1, \quad (1.5)$$

where $v = (\sinh^{-1} 1)/t$. This will be shown below in the general case. These lemniscates are the same as the 1-lemniscates of $T_t(z')$. The w -plane, where $w = u + iv = \cos(z')$, will be useful later. The lemniscates in different domains are shown in Figure III-1.

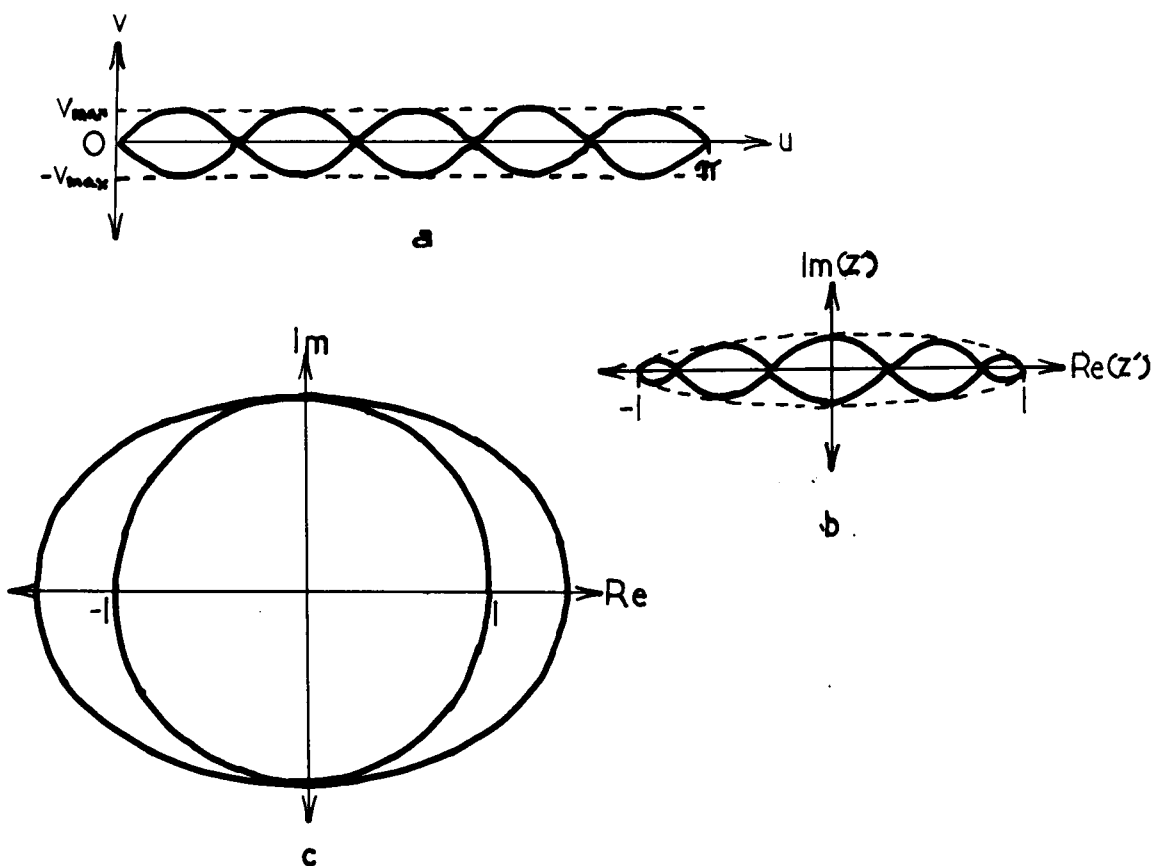


Figure III-1. The 1-lemniscates of $T_t(z')$, with $t=5$, under different transformations.

- a) The region $w = u+iv$ such that $0 \leq u \leq \pi$ and $|T_t(\cos w)| \leq 1$. The boundary oscillates between $v = v_{\min} = 0$ and $|v| = v_{\max} = (\sinh^{-1} 1)/t$.
- b) The region $z' = \cos(u+iv)$ such that $|T_t(z')| \leq 1$. The boundary oscillates between the degenerate ellipse which is the real line $[0,1]$ and the ellipse of equation (1.5).
- c) The circle $|T_t(z')| \leq 1$, the degenerate ellipse which is the real line $[-1,1]$, and the ellipse $\left[\frac{\operatorname{Re}(z')}{\cosh tv_{\max}} \right]^2 + \left[\frac{\operatorname{Im}(z')}{\cosh tv_{\max}} \right]^2 = 1$. As u varies from 0 to 2π with $|T_t(z')| = 1$, $T_t(z')$ traces the circle t times and $|v|$ oscillates between 0 and $v_{\max}$ as shown in a) for $0 \leq u \leq \pi$.

Other larger regions for which the optimal t -degree polynomial normalized at $z' = \mu$ is $P_t(z') = T_t(z')/T_t(\mu)$ can be found by considering the $\gamma/T_t(\mu)$ -lemniscates of these polynomials when $1/T_t(\mu) \leq \gamma/T_t(\mu) \leq 1$. We have the following result.

Theorem 1: The Chebyshev polynomials $P_t(z') = T_t(z')/T_t(\mu)$ are optimal on their $\gamma/T_t(\mu)$ -lemniscates, where $1 \leq \gamma \leq T_t(\mu)$, over the t -degree polynomials normalized at $z' = \mu$.

Proof: Suppose there exists a t -degree polynomial $Q_t(z')$ with $Q_t(\mu) \doteq 1$ and $|Q_t(z')| \leq H < \gamma/T_t(\mu)$ on the boundary of L , $L = \{z': |P_t(z')| \leq \gamma/T_t(\mu)\}$. $P_t(z')$ has t roots in L . Therefore by Rouché's Theorem (see Appendix A) $R_t(z') = P_t(z') - Q_t(z')$ also has t roots in L . But $R_t(\mu) = 1 - 1 = 0$, and μ is not an element of L , so $R_t(z')$ has at least $t+1$ roots, which is a contradiction since $R_t(z')$ is a t -degree polynomial. $\square$

In the z' -plane, the $\gamma/T_t(\mu)$ -lemniscate of $P_t(z')$ is the same as the γ -lemniscate of $T_t(z')$. To describe it, let $w = u+iv = \cos^{-1}(z')$, so that

$$\begin{aligned} 1 < \gamma^2 &= |T_t(z')|^2 = |\cos(t \cos^{-1} z')|^2 = |\cos t(u+iv)|^2 \\ &= \cos^2(tu) \cosh^2(tv) + \sin^2(tu) \sinh^2(tv). \end{aligned} \quad (1.6)$$

Therefore as u varies from 0 to π , v oscillates between

$v_{\min} = (\cosh^{-1} \gamma)/t$ and $v_{\max} = (\sinh^{-1} \gamma)/t$, and $z' = \cos(u+iv)$. The γ -lemniscate of $T_t(z')$ in the z' -plane oscillates between the ellipses

$$(\operatorname{Re}(z')/\cosh v_{\min})^2 + (\operatorname{Im}(z')/\sinh v_{\min})^2 = 1 \quad (1.7)$$

and

$$(\operatorname{Re}(z')/\cosh v_{\max})^2 + (\operatorname{Im}(z')/\sinh v_{\max})^2 = 1, \quad (1.8)$$

as shown in Figure III-2.

We approximate this lemniscate by the largest ellipse it contains, E ,

$$E = \{z': (\operatorname{Re}(z')/\cosh v_{\min})^2 + (\operatorname{Im}(z')/\sinh v_{\min})^2 \leq 1\}. \quad (1.9)$$

It has been shown that the polynomial $P_t(z')$ is in fact the optimal polynomial on this ellipse (Manteuffel [1977]). The approximation of the lemniscate by the ellipse is justified also by considering the lemniscate as t gets large when $v_{\min}$ is fixed, that is, when the lemniscate must contain certain eigenvalues. Consider the points on the boundary of L which are strictly real or strictly imaginary. Then the condition $|T_t(z')| \leq \gamma$ implies $\sinh tv_{\max} = \cosh tv_{\min}$, which implies that $v_{\max}$ approaches $v_{\min}$ as t gets large, therefore the lemniscate approaches the ellipse. From this it is evident that as t gets large the image of the ellipse under the Chebyshev function T_t approaches the image of the lemniscate under T_t , which is of course a circle. The significance of this is addressed by Trefethen [1981].

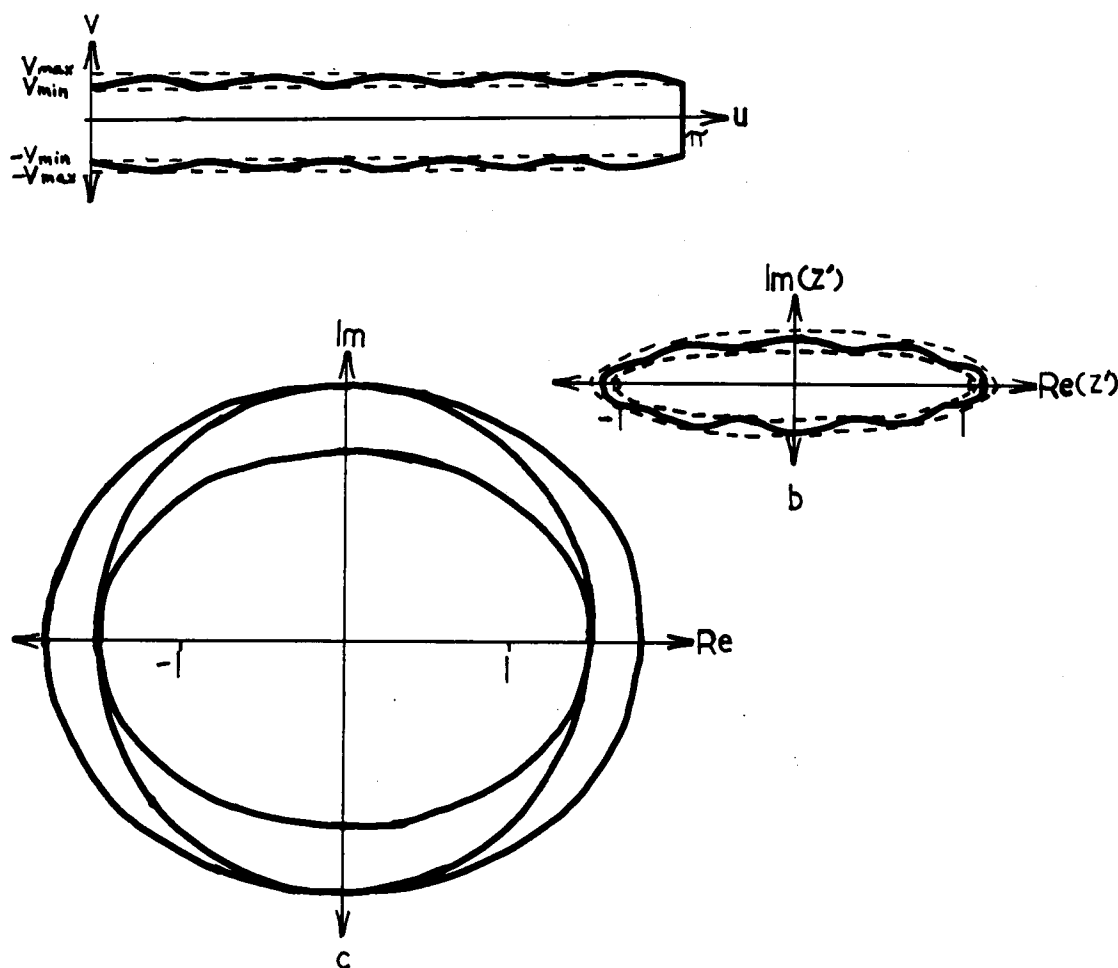


Figure III-2. The γ -lemniscates of $T_t(z')$, with $t=5$ and $\gamma=1.5$, under the same transformations as in Figure 1.1.

- The region $w = u+iv$ such that $0 \leq u \leq \pi$ and $|T_t(\cos w)| \leq \gamma$. The boundary oscillates between $|v| = v_{\min} = (\cosh^{-1} \gamma)/t$ and $|v| = v_{\max} = (\sinh^{-1} \gamma)/t$.
- The region $z' = \cos(u+iv)$ such that $|T_t(z')| \leq \gamma$. The boundary oscillates between the ellipse in equation (1.7) and that in equation (1.8).
- The circle $|T_t(z')| \leq \gamma$, the ellipse $\cos(u+iv_{\min})$ and the ellipse $\cos(u+iv_{\max})$. As u varies from 0 to 2π with $|T_t(z')| = \gamma$, $T_t(z')$ traces the circle t times and $|v|$ oscillates between $v_{\min}$ and $v_{\max}$, as shown in a) for $0 \leq u \leq \pi$.

Briefly, a best approximation to a function on a region, in this case the function zero on the ellipse, will have an error curve on the boundary which approaches circularity and has decreasing radius as the degree of the approximating function gets large.

To apply this theory, let the original eigenvalue spectrum be transformed to lie in an ellipse in the z -plane with major axis on the real axis between a and b , $a < b < 1$, and minor axis length $2c$. Then the transformation $z' = (\cosh v_{\min}) \times \left[\frac{2z - (a+b)}{b-a} \right]$, where $v_{\min} = \tanh^{-1} \left[\frac{2c}{b-a} \right]$, takes this ellipse to the ellipse E . Therefore the optimal polynomial is

$$P_t(z') = T_t(z') / T_t(\mu \cosh v_{\min}). \quad (1.10)$$

The error reduction for t Chebyshev iterations is

$$\begin{aligned} \max_{z' \in E} |P_t(z')| &= \frac{T_t(\cosh v_{\min})}{T_t(\mu \cosh v_{\min})} \\ &= \frac{\frac{1}{2} [\tan^t \frac{\phi}{2} + \tan^{-t} \frac{\phi}{2}]}{\cosh [t \cosh^{-1} \frac{\mu}{\sin \phi}]} \end{aligned} \quad (1.11)$$

where $\cos \phi = \frac{2c}{b-a}$ (Wachspress, personal communication).

2. Lemniscates and Their Approximating Ellipses

For Analysis of ADI Iteration

In ADI iteration for a real spectrum in an interval $[k', 1]$, the optimum iteration parameters ω_j are those which minimize the maximum absolute value of

$$Q_t(x, \omega) = \prod_{j=1}^t \frac{(x - \omega_j)}{(x + \omega_j)} \quad 0 < k' \leq x \leq 1, \quad (2.1)$$

or substituting the elliptic function $\text{dn}(Ku, k) = x$,

$$Q_t(u) = \prod_{j=1}^t \frac{\text{dn}(Ku, k) - \omega_j}{\text{dn}(Ku, k) + \omega_j} \quad 0 \leq u \leq 1. \quad (2.2)$$

The optimum ω_j will be those which make this function equal to

$$Q_t(u) = \frac{\text{dn}(tK_1 u, k_1) - \sqrt{k_1}}{\text{dn}(tK_1 u, k_1) + \sqrt{k_1}} \\ = \sqrt{k_2} \text{sn}[K_2(2tu+1), k_2] \quad 0 \leq u \leq 1 \quad (2.3)$$

where k_1 is determined by $K'_1/K_1 = tK'/K$ and $\sqrt{k_2} = (1 - \sqrt{k_1})/(1 + \sqrt{k_1})$. Then $\omega_j = \text{dn}((2j-1)K/2t, k)$, $j=1, 2, \dots, t$ (analysis of W.B. Jordan, see Wachspress [1966]). Since the function is optimal for $x = \text{dn}(Ku, k)$, $0 \leq u \leq 1$, it is also optimal for complex $z = \text{dn}(Kw, k)$, $w = u+iv$, where w is determined by $|\text{sn}[K_2(2tw+1), k_2]| \leq 1$ (Wachspress [1966]).

Some notation will be useful later. The original eigenvalue spectrum is transformed by W.B. Jordan's bilinear transformation (Wachspress [1966]) to lie in a region Λ of the complex plane where $k' \leq \text{Re}(\lambda) \leq 1$. The region W of the complex plane consists of those $w = u+iv$ where $\lambda = \text{dn}(Kw, k) \in \Lambda$ when $w \in W$. Another useful region is the 'ln-symmetric' region Z' where $z' \in Z'$ when $z' = \ln[(1/\sqrt{k'})\lambda] = \ln[(1/\sqrt{k'}) \text{dn}(Kw, k)]$ for $\lambda \in \Lambda$ and,

equivalently, $w \in W$. Figure III-3 shows the regions $|\operatorname{sn}[K_2(2tw+1), k_2]| \leq 1$ under these transformations.

The ADI theory is extended to the case of complex eigenvalues in the same manner as Chebyshev extrapolation theory. The function $\sqrt{k_2}[\operatorname{sn}(K_2(2tw+1), k_2)]$, $0 \leq \operatorname{Re}(w) \leq 1$, yields the optimum parameters described above when transformed eigenvalues lie in its $\sqrt{k_2}$ -lemniscate, which is the same as the 1-lemniscate of $\operatorname{sn}[K_2(2tw+1), k_2]$. Therefore consider the $\gamma\sqrt{k_2}$ -lemniscate, $1 \leq \gamma \leq 1/\sqrt{k_2}$, with $\operatorname{Re}(w) = u$ still restricted to $[0, 1]$ (Figure III-4). We have the following theorem.

Theorem 2: The functions $Q_t(z, \omega) = \prod_{j=1}^t \frac{(z - \omega_j)}{(z + \omega_j)}$, where ω_j are the optimum ω_j determined above, are optimal on their $\gamma\sqrt{k_2}$ -lemniscates, where $1 \leq \gamma \leq 1/\sqrt{k_2}$, over all rational functions of type $\frac{p_t(-z)}{p_t(z)}$, where $p_t(z)$ is a t -degree polynomial in z .

Proof: Let $q_t(z) = \prod_{j=1}^t (z + \omega_j)$. Then $q_t(-z) = \prod_{j=1}^t (-z + \omega_j)$ and $Q_t(z) = \frac{q_t(-z)}{q_t(z)}(-1)^t$. Suppose there exists a polynomial $P_t(z) = \frac{p_t(-z)}{p_t(z)}$ such that for all z on the boundary of L , $|P_t(z)| < |Q_t(z)|$. The difference between these functions, $R = Q_t - (-1)^t P_t = \frac{q_t(-z)p_t(z) - q_t(z)p_t(-z)}{p_t(z)q_t(z)} = \frac{N_{2t}(z)}{D_{2t}(z)}$, has

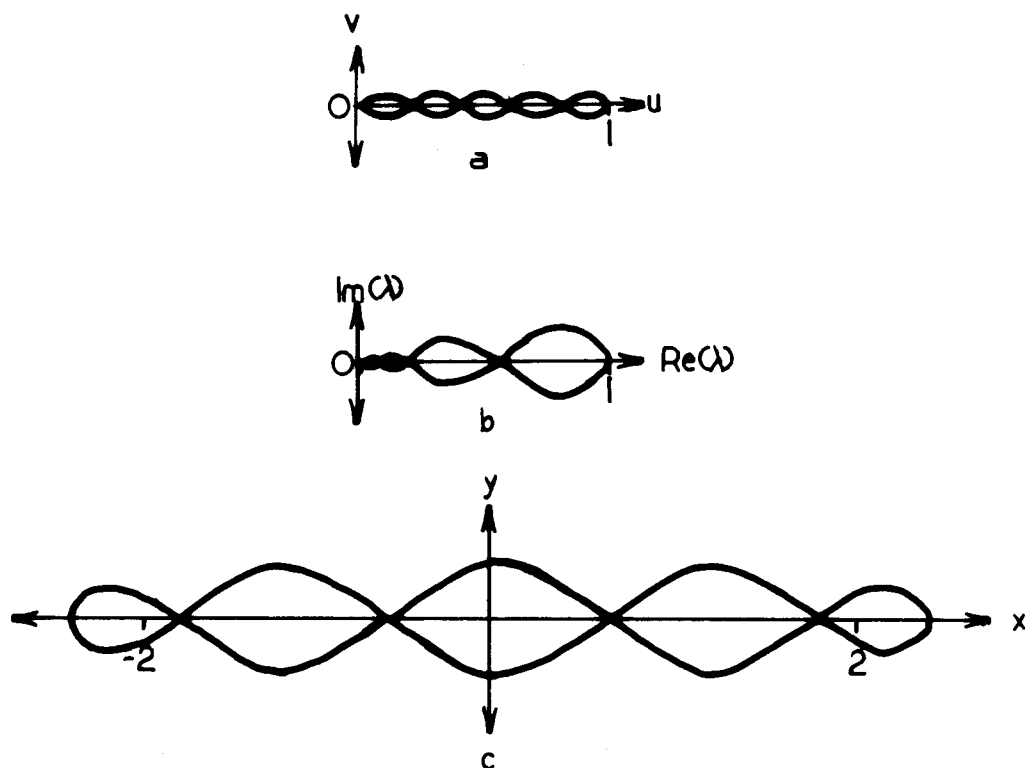


Figure III-3. The 1-lemniscates of $\text{sn}[K_2(2tw+1), k_2]$, with $t=5$, under different transformations.

- a) The region W of $w = u+iv$ such that $0 \leq u \leq 1$ and $|\text{sn}[K_2(2tw+1), k_2]| \leq 1$.
- b) The region Λ of $\lambda = \text{dn}(Kw, k)$ where $\lambda \in \Lambda$ when $W \in w$. The real extent of this region is from k' to 1. Note that if λ is in the region then $1/\lambda$ is in the region.
- c) The region Z' of $z' = x+iy = \ln \left[\frac{1}{\sqrt{k'}} \lambda \right]$
 $= \ln \left[\frac{1}{\sqrt{k'}} \text{dn}(Kw, k) \right]$ where $z' \in Z'$ when $\lambda \in \Lambda$ and, equivalently, $w \in W$.

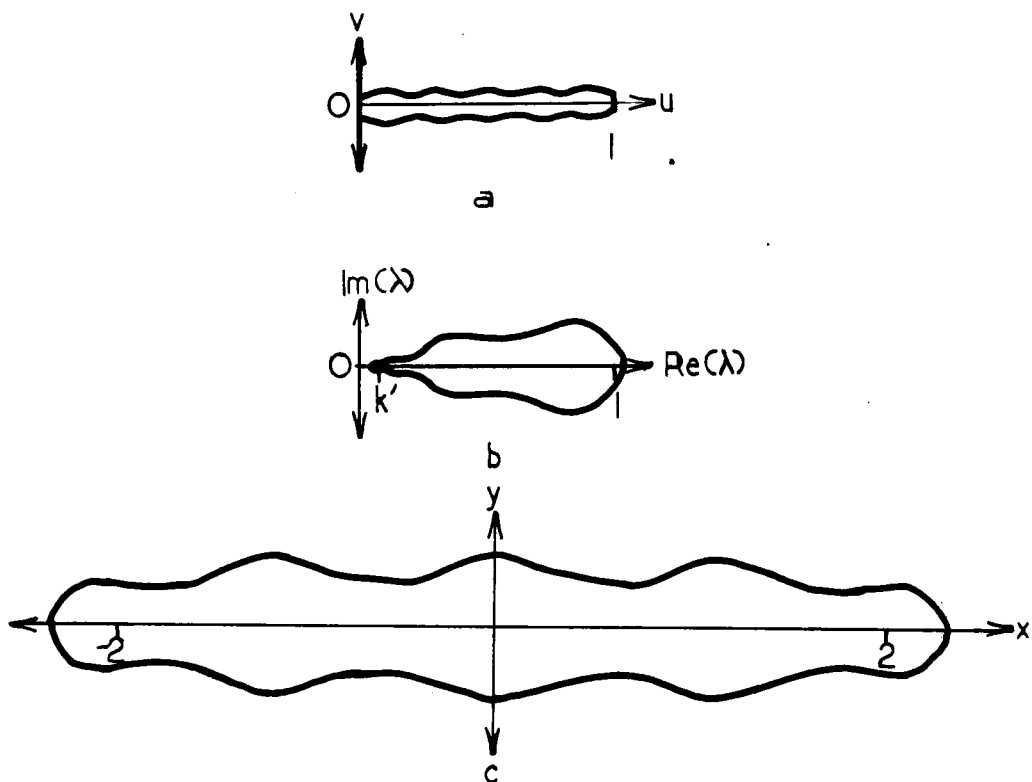


Figure III-4. The γ -lemniscates of $\text{sn}[K_2(2tw+1), k_2]$, with $t=5$ and $\gamma=1.5$, under the same transformations as in Figure III-3.

- a) The region W of $w = u+iv$ such that $0 \leq u \leq 1$ and $|\text{sn}[K_2(2tw+1), k_2]| \leq \gamma$.
- b) The region Λ of $\lambda = \text{dn}(Kw, k)$ where $\lambda \in \Lambda$ when $w \in W$.
- c) The region Z' of $z' = x+iy = \ln\left(\frac{1}{\sqrt{k'}} \lambda\right)$ where $z' \in Z'$ when $\lambda \in \Lambda$ and, equivalently, $w \in W$.

maximal degree $2t-1$. By Rouché's Theorem (see Appendix A), R has t roots in L . Also, $N_{2t}(-z) = -N_{2t}(z)$ so R has t roots in the mirror image of L , $\{z = x-iy: -x-iy \in L\}$, and $N_{2t}(0) = 0$. Thus R has $2t+1$ roots, which is a contradiction. □

We thus have optimum parameters for eigenvalue regions which, when appropriately transformed and ln-symmetrized, lie in the $\gamma\sqrt{k_2}$ -lemniscate of the sn function above. The ADI spectral radius for t iterations is $\gamma^2 k_2$.

The largest convex region in a lemniscate can be considered as a conservative bound for eigenvalues within it. That is, if the spectrum lies in this conservative region it lies in the associated lemniscate. The conservative region is of interest because it can be fixed while t varies and is approximately elliptical, as will be seen later. The $\gamma\sqrt{k_2}$ -lemniscate in the w -plane oscillates between $v = v_{\min}$ and $v = v_{\max}$, so the conservative region is $W_c = \{w = u+iv: 0 \leq u \leq 1, 0 \leq v \leq v_{\min}\}$. This maps into the conservative regions Λ_c and Z'_c , shown in Figure III-5.

The lemniscate is an improving approximation to the conservative region as t gets large. This property is essential for the rational function to be optimal on the conservative region, just as in the case of Chebyshev polynomial approximation on an ellipse (Trefethen [1981]). To see that this property holds, consider the following. In

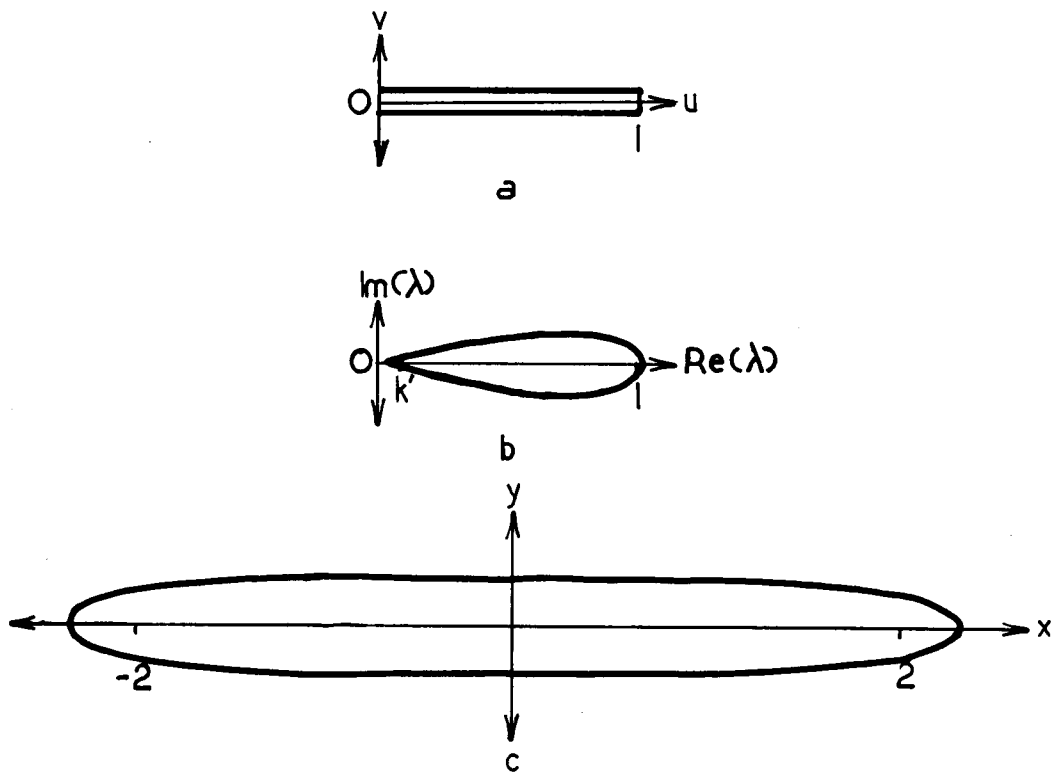


Figure III-5. The conservative regions of the lemniscates shown in Figure III-4.

- a) W_c
- b) $\Lambda_c = \text{dn}(W_c K, k)$
- c) $Z'_c = \ln\left(\frac{1}{\sqrt{k'}} \Lambda_c\right)$

order for the transformed eigenvalues to remain in the lemniscate as t increases, they must be in the conservative region Z'_c because the scallops outside Z'_c are shifting as t changes. Fixing Z'_c so it doesn't decrease in size as t increases is equivalent to fixing k' and $v_{\min}$. Then k_2 , already assumed small, decreases with increasing t so we may approximate for t large

$$\begin{aligned} \operatorname{sn}[(1+i2tv_{\min})K_2, k_2] &= 1/\operatorname{dn}(2tv_{\min} K_2, k_2') \\ &= \cosh(\pi tv_{\min}) \\ &= \gamma \end{aligned} \quad (2.4)$$

and

$$\sinh(\pi tv_{\max}) = \gamma. \quad (2.5)$$

Thus for t increasingly large, $\gamma = \sinh(\pi tv_{\max}) = \cosh(\pi tv_{\min})$ is satisfied by $v_{\max}$ increasingly close to $v_{\min}$, and the lemniscate approaches the conservative region Z'_c .

The region Z'_c can be approximated by an ellipse. We will use this to approximate the error reduction when the transformed eigenvalues lie in an ellipse. The ellipse with the smallest possible value of $\gamma\sqrt{k_2}$ determines the optimum parameters, which give error reduction $\gamma\sqrt{k_2}$. To show that Z'_c can be approximated by the ellipse E ,

$$E = \{x+iy: \left[\frac{x}{\ln(\sqrt{k'} \cos(v_{\min} \ln \frac{4}{k'}))} \right]^2 + \left[\frac{y}{v_{\min} \ln \frac{4}{k'}} \right]^2 \leq 1\}, \quad (2.6)$$

note that Z'_c is symmetric about the real and imaginary axes

and always has a somewhat elliptical shape. Then the approximation follows from two calculations.

The first calculation shows that $y_{\max} = \text{maximum } |y|$ occurs when $x=0$, and $y_{\max} \cong v_{\min} \ln(4/k')$. This is because $x+iy = z = \ln[(1/\sqrt{k'}) \text{dn}(Ku+iKv_{\min},k)]$ so that

$$\begin{aligned} e^x e^{iy} &= \frac{1}{\sqrt{k'}} \text{dn}(Ku+iKv_{\min},k) \\ &= \frac{1}{k'} \left[\frac{\text{dn}(Ku,k) \text{cn}(Kv_{\min},k') \text{dn}(Kv_{\min},k')}{\text{cn}^2(Kv_{\min},k') + k^2 \text{sn}^2(Ku,k) \text{sn}^2(Kv_{\min},k')} \right. \\ &\quad \left. - \frac{ik^2 \text{sn}(Ku,k) \text{cn}(Ku,k) \text{sn}(Kv_{\min},k')}{\text{cn}^2(Kv_{\min},k') + k^2 \text{sn}^2(Ku,k) \text{sn}^2(Kv_{\min},k')} \right] \quad (2.7) \end{aligned}$$

and therefore

$$y = \tan^{-1} \left[\frac{k^2 \text{sn}(Kv_{\min},k')}{\text{cn}(Kv_{\min},k') \text{dn}(Kv_{\min},k')} \cdot \frac{\text{sn}(Ku,k) \text{cn}(Ku,k)}{\text{dn}(Ku,k)} \right] \quad (2.8)$$

where the first quotient is fixed so that $y_{\max}$ occurs at the maximum of the second quotient. To find it, set the derivative of the second quotient equal to zero, that is

$$\begin{aligned} \frac{d}{du} \left[\frac{\text{sn}(Ku,k) \text{cn}(Ku,k)}{\text{dn}(Ku,k)} \right] \\ = \frac{K}{\text{dn}^2(Ku,k)} \text{dn}^2(Ku,k) \text{cn}^2(Ku,k) - \text{dn}^2(Ku,k) \text{sn}^2(Ku,k) \\ + k^2 \text{sn}^2(Ku,k) \text{cn}^2(Ku,k) = 0 \quad (2.9) \end{aligned}$$

so that

$$\text{cn}^2(Ku,k) [\text{dn}^2(Ku,k) + k^2 \text{sn}^2(Ku,k)] - \text{sn}^2(Ku,k) \text{dn}^2(Ku,k) = 0 \quad (2.10)$$

and therefore

$$\text{cn}(Ku, k) = \text{sn}(Ku, k) \text{dn}(Ku, k). \quad (2.11)$$

This occurs at exactly one point for $0 \leq u \leq 1$, when $u = \frac{1}{2}$. We have

$$\text{sn}(K/2, k) \text{dn}(K/2, k) = \frac{\sqrt{k'}}{\sqrt{1+k'}} = \text{cn}(K/2, k), \quad (2.12)$$

so that

$$\begin{aligned} y_{\max} &= \tan^{-1} \left[\frac{k^2 \text{sn}(Kv_{\min}, k')}{\text{cn}(Kv_{\min}, k') \text{dn}(Kv_{\min}, k')} \cdot \frac{1}{1+k'} \right] \\ &\cong \tan^{-1} \left[\frac{\sin Kv_{\min}}{\cos Kv_{\min}} (1-k') \right] \\ &\cong Kv_{\min} \\ &\cong \left(\ln \frac{4}{\sqrt{k'}} \right) v_{\min}. \end{aligned} \quad (2.13)$$

Now, when $u = \frac{1}{2}$, $x = 0$ because $\left| \frac{1}{\sqrt{k'}} \text{dn}\left(\frac{K}{2} + iKv_{\min}, k\right) \right| = 1$ (since the line $w = \frac{K}{2} + iv$ transforms to the circle $|z| = |\text{dn } w| = \sqrt{k'}$). The second calculation shows that $x_{\max} = \text{maximum } |x| \cong \ln \left[\sqrt{k'} \cos(v_{\min} \ln \frac{4}{\sqrt{k'}}) \right]$. To show this, note that the maximum of $\left| \frac{1}{\sqrt{k'}} \text{dn}[K(u+iv_{\min}), k] \right|$ for $0 \leq u \leq 1$ will occur at $u=0$ and $u=1$, and therefore will be real so that $y=0$. Also, the maximum value is approximately the inverse of the minimum value, so at $u=0$,

$$\begin{aligned} \ln \left| \frac{1}{\sqrt{k'}} \text{dn}(ikv_{\min}, k) \right| &= \ln \left| \frac{1}{\sqrt{k'}} \frac{\text{dn}(Kv_{\min}, k')}{\text{cn}(Kv_{\min}, k')} \right| \\ &\cong -\ln(\sqrt{k'} \cos Kv_{\min}). \end{aligned} \quad (2.14)$$

3. An Estimate of the Effective Spectral Radius Of ADI Iteration

It has been shown that in the case of real eigenvalues the error reduction after t ADI steps with optimum parameters is approximated well by

$$\mu(t) = 4e^{\pi^2 t / \ln(k'/4)} \quad (3.1)$$

so the effective spectral radius is

$$\mu = \lim_{t \rightarrow \infty} [\mu(t)]^{1/t} = e^{\pi^2 / \ln(k'/4)} \quad (3.2)$$

(Wachspress [1966]). A derivation with properties of elliptic functions gives an estimate of the effective spectral radius for ADI iteration with complex spectra as a function of k' and $v_{\min}$.

We will need the following lemma.

Lemma: For $0 \leq s \leq \frac{1}{2}$,

$$1 \leq \frac{1}{\operatorname{dn}(sK, k)} \leq (k')^{-s} \leq 1 + 2[(k')^{-\frac{1}{2}} - 1]s. \quad (3.3)$$

Proof: The equality holds at $s=0$ and $s=\frac{1}{2}$ since $\operatorname{dn}(0, k) = 1$ and $\operatorname{dn}(K/2, k) = \sqrt{k'}$. In between, dn is concave down while $(k')^s$ is concave up because $\frac{d^2}{ds^2}(k')^s = (\ln k')^2 (k')^s \geq 0$. Therefore $\operatorname{dn}(sK, k) \geq (k')^s$ and $1/\operatorname{dn}(sK, k) \leq (k')^{-s}$. It is also true that $(k')^{-s}$ is concave up, hence the second inequality. □

A standard approximation for the nome q_1 is

$$q_1 = \epsilon_1 + 2\epsilon_1^5 + 15\epsilon_1^9 + \dots, \quad \epsilon_1 = \frac{1 - \sqrt{k'_1}}{2(1 + \sqrt{k'_1})}. \quad (3.4)$$

Since $\epsilon_1 = \sqrt{k_2}/2$ we have the nome in terms of k_2

$$q_1 = \frac{\sqrt{k_2}}{2} + 2\left[\frac{\sqrt{k_2}}{2}\right]^5 + \dots \quad (3.5)$$

When k_2 is small, solving for k_2 and using analysis from the real eigenvalue case of ADI iteration yields

$$\sqrt{k_2} \cong 2q_1 = 2q^t = 2e^{\pi^2 t / 2 \ln(k'/4)}. \quad (3.6)$$

Substituting this in a well-known approximation for K'/K gives

$$\frac{K'_2}{K_2} = \frac{2}{\pi} \ln \left[\frac{4}{k_2} \right] = \frac{2\pi t}{\ln(4/k')}. \quad (3.7)$$

In order to obtain an estimate of $\max |Q_t|^2$ in terms of k' and $v_{\min}$ we apply the lemma so that

$$\begin{aligned} \mu(t) &= \max_{0 \leq u \leq 1} \left[\sqrt{k_2} \operatorname{sn}[K_2(2tw+1), k_2] \right]^2 \\ &= \left[\sqrt{k_2} / \operatorname{dn}(2v_{\min} t K_2, k'_2) \right]^2 \\ &\leq (k_2)^{1-4tv_{\min}(K_2/K'_2)} \\ &= 4^{\frac{2v_{\min} \ln(4/k')}{\pi}} e^{t(\pi^2 / \ln(k'/4) + 2\pi v_{\min})} \end{aligned} \quad (3.8)$$

The effective spectral radius is then

$$\lim_{t \rightarrow \infty} [\mu(t)]^{1/t} = e^{(\pi^2 / \ln(k'/4) + 2\pi v_{\min})} \quad (3.9)$$

which reduces to the solution in the case of real eigenvalues when $v_{\min} = 0$.

The upper limit on $v_{\min}$ for convergence of the method is consistent with the limits on s in the lemma, with both requiring $v_{\min} < \pi / 2 \ln(k'/4)$. The same bound is obtained by requiring $v_{\min} < v_0$ where $\operatorname{dn}(2tv_0 K_2, k_2) = \sqrt{k_2}$, that is $2tv_0 K_2 = K_2'/2$ or $v_{\min} < K_2' / 4tK_2$. Clearly, the elliptic functions stay well away from any of their poles in the complex plane regardless of t .

CHAPTER IV

COMPUTER IMPLEMENTATION AND RESULTS

1. Methods

In the previous chapter the optimum parameters for ADI iteration were found to be computed from the $v_{\min}$ and k' which maximize $2F(k', v_{\min}) = \pi/\ln(4/k') - 2v_{\min}$ subject to the condition $\ln(\frac{1}{\sqrt{k'}} \lambda_i) \in E$, where λ_i are the eigenvalues when transformed so that $\text{Re}(\lambda_i)$ is in $[k', 1]$. The optimum parameters are then $\omega_j = \text{dn}\{[(2j-1)/2t]K(k_{\text{opt}}), k_{\text{opt}}\}$ where k_{opt} is the optimum value of k (see Wachspress [1966]). This section describes a method of finding the optimum value of k , and therefore the optimum parameters ω_j , when the spectrum of the iteration matrix lies in a convex polygon with positive real components and comparatively small complex components. The results of some tests using this method and their relation to Wachspress' theory for eigenvalues on a circle are discussed in the remainder of the chapter.

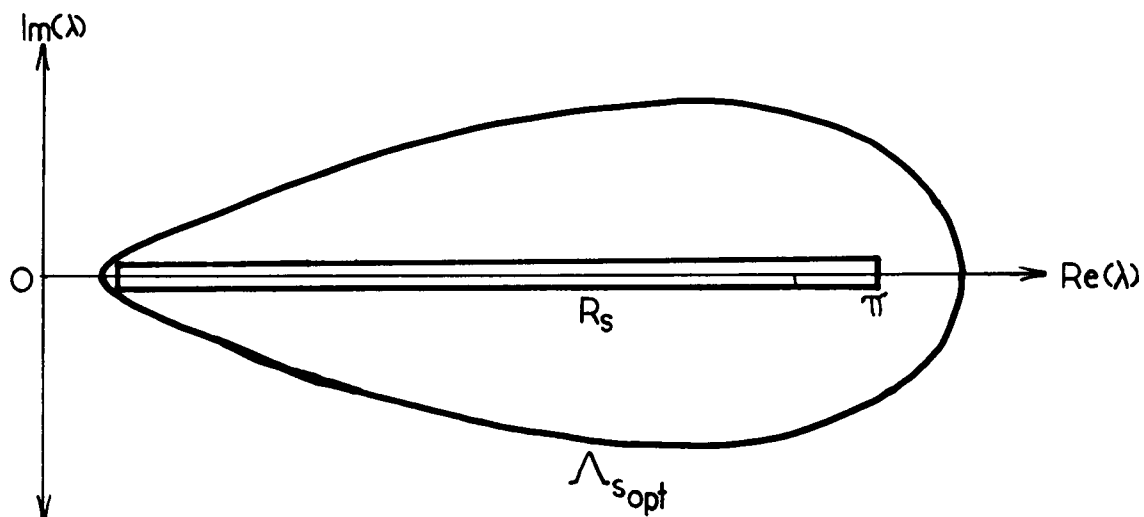
First, consider the case where the spectrum has strictly positive real part and is described by maximum and minimum real bounds and by a maximum complex magnitude bound, that is the spectrum lies in a rectangle. The spectrum will always be symmetric about the real axis. We can perform a scaling on this region by multiplying it by a

factor determined by the coordinates of the original rectangle so that the real part of the new rectangle R_s lies between some a in $(0,1)$ and $1/a$, and the complex part lies between $-ib$ and ib , b positive. To find optimum iteration parameters, this rectangle must lie in the region $\frac{1}{\sqrt{k'}} \Lambda_c = \Lambda_{s_{opt}} = \text{dn}(W_c K, k) / \sqrt{k'}$ with maximum $F(k', v_{min}) = \pi/2 \ln(4/k') - v_{min}$ of all such regions Λ_s containing the rectangle R_s (see Figure IV-1). Clearly, a Λ_s region which may maximize F will include one of the corners of the rectangle R_s on its boundary, as otherwise Λ_s could be shrunk and F increased by decreasing v_{min} or increasing k' until that happened. This remains true when R_s and Λ_s are transformed by taking natural logarithms of the regions.

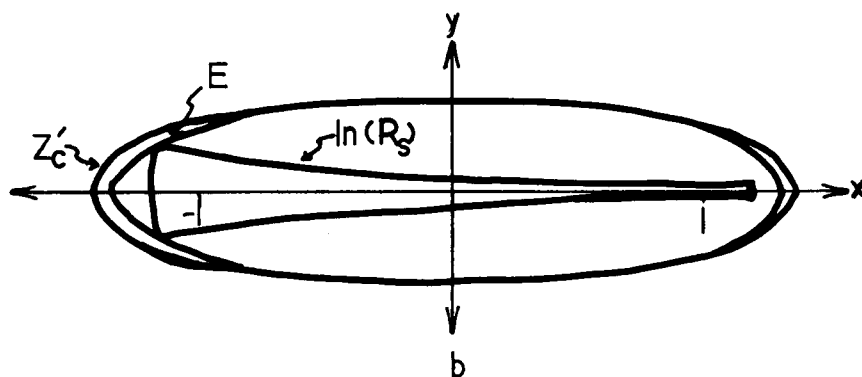
In order to facilitate the calculations, the region $Z'_c = \ln(\Lambda_s)$ is approximated by the ellipse E of Chapter 3. The problem then becomes the maximization of $F(k', v_{min})$ over k' and v_{min} when constrained so that the boundary of E , which is a function of k' and v_{min} , contains at least one of the corners of $\ln(R_s)$. This constraint can be written as

$$\frac{A^2}{[\ln(\sqrt{k'} \cos(v_{min} \ln \frac{4}{k'}))]^2} + \frac{B^2}{[v_{min} \ln \frac{4}{k'}]^2} = 1 \quad (1.1)$$

where A and B depend on a and b in a manner which will be discussed shortly. This problem can be expressed as a constrained maximization problem in one variable by substituting $\eta = v_{min} \ln(4/k')$. Then η is chosen to maximize



a



b

Figure IV-1. A rectangular spectrum R_s normalized so its real extent is $\ln$ -symmetric, shown with the associated region $\Lambda_{s_{opt}} = \text{dn}(W_c K, k) / \sqrt{k'}$, for some k and v_{min} where $\text{Im}(w) \leq v_{min}$, computed from the ellipse E with fastest theoretical convergence of those E where $\ln(R_s) \subset E$.

a) R_s and $\Lambda_{s_{opt}}$.

b) $\ln(R_s)$, $Z'_c = \ln(\Lambda_{s_{opt}})$, and E .

$$F(\eta) = (\frac{\pi}{2} - \eta) / \left[2 \left[\frac{A\eta}{\sqrt{\eta^2 - B^2}} + \ln(\cos \eta) + \ln 2 \right] \right], \quad (1.2)$$

and k' and $v_{\min}$ are found as functions of η by

$$k'(\eta) = e^{-2[(A\eta/\sqrt{\eta^2 - B^2}) + \ln(\cos \eta)]} \quad (1.3)$$

and

$$v_{\min}(\eta) = \eta/2 \left[\frac{A\eta}{\sqrt{\eta^2 - B^2}} + \ln(\cos \eta) + \ln 2 \right]. \quad (1.4)$$

The limits on η for these expressions to make sense are $B < \eta < \eta_{\sup}(A,B) < \pi/2$ where $\eta_{\sup}$ is the smallest η greater than B such that

$$\ln(\cos \eta_{\sup}) + \frac{A\eta_{\sup}}{\sqrt{\eta_{\sup}^2 - B^2}} = 0. \quad (1.5)$$

Optimum iteration parameters in terms of the solution η_{opt} of the problem above are $\omega_j = \text{dn}\{[(2j-1)/2t]K(k(\eta_{\text{opt}})), k(\eta_{\text{opt}})\}$. The error reduction factor is $k_2/\text{dn}(2tv_{\min}K_2, k_2')$.

The variables A and B are the absolute real and imaginary components of the corner of $\ln(R_s)$ that lies on the boundary of the ellipse. Since η is one half the minor axis of the ellipse, which lies along the imaginary axis with midpoint at the origin, it is clear that the lower limit on η is B where B is the imaginary component of the corner of $\ln(R_s)$ which has largest imaginary component. A is the real component of this corner. Increasing η from its

lower limit increases the minor axis and decreases the major axis of the ellipse defined by this A, B, and η above. Eventually, η reaches some value η_c where the ellipse contains all corners of $\ln(R_s)$ (recall that all regions under discussion are symmetric about the real axis). In order to continue increasing η while maintaining $\ln(R_s)$ within the ellipse it is necessary to change the values of A and B to be the absolute values of the coordinates of the other corner (see Figure IV-2). Thus $F(\eta)$ is fully defined for a spectrum bounded by a rectangle when A and B are the functions of η_c defined by

A = real component of the complex number whose
imaginary component is B

$$B = \max(|\operatorname{Im} \ln(a+ib)|, |\operatorname{Im} \ln(1/a + ib)|) \quad \eta \leq \eta_c$$

$$\min(|\operatorname{Im} \ln(a+ib)|, |\operatorname{Im} \ln(1/a + ib)|) \quad \eta \geq \eta_c.$$

The value of η_c is easy to find since it satisfies

$$\frac{[\operatorname{Re} \ln(a+ib)]^2}{c\eta_c^2} + \frac{[\operatorname{Im} \ln(a+ib)]^2}{\eta_c^2} = 1 \quad \text{where} \quad (1.6)$$

$$c = \left| \frac{[\operatorname{Re} \ln(a+ib)]^2 - [\operatorname{Re} \ln(\frac{1}{a}+ib)]^2}{[\operatorname{Im} \ln(a+ib)]^2 - [\operatorname{Im} \ln(\frac{1}{a}+ib)]^2} \right|$$

or

$$\eta_c = \sqrt{\frac{[\operatorname{Re} \ln(a+ib)]^2}{c} + [\operatorname{Im} \ln(a+ib)]^2}. \quad (1.7)$$

The function $F(\eta)$ is continuous but not necessarily differentiable at η_c .

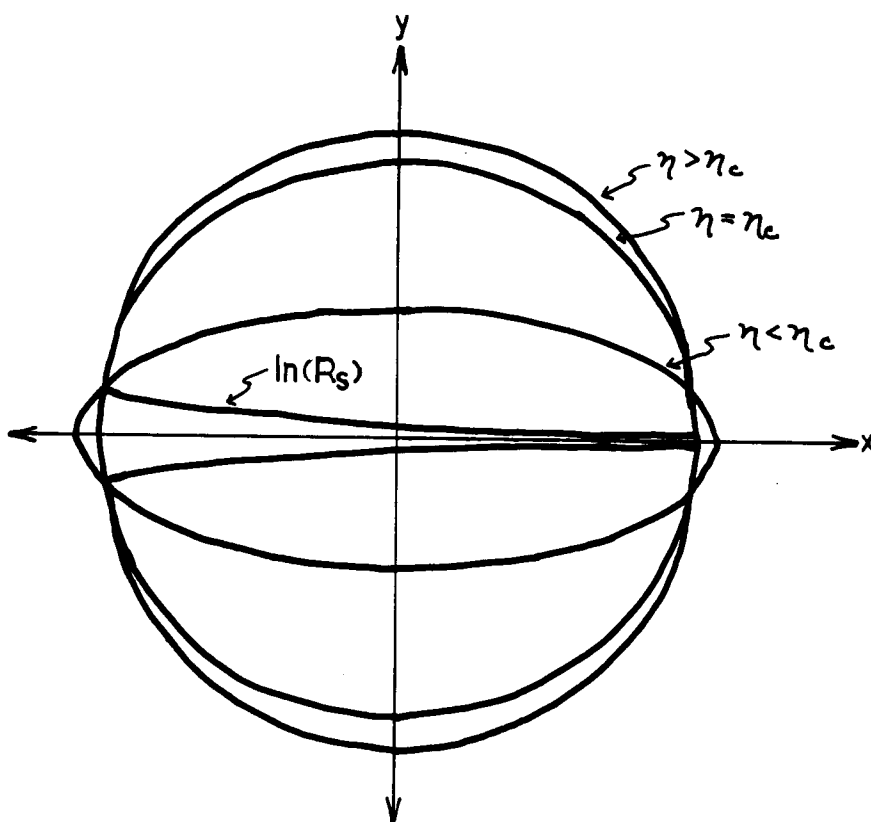


Figure IV-2. Natural logarithm of a $\ln$ -symmetrized rectangular spectrum R_s with three sample regions E . The regions E shown are for $\eta = \eta_c$ where the boundary of E contains the four corners of $\ln(R_s)$, a value $\eta < \eta_c$ where the boundary of E contains the lefthand corners of $\ln(R_s)$, and a value $\eta > \eta_c$ where the boundary of E contains the righthand corners of $\ln(R_s)$.

These ideas extend easily to spectra bounded by convex polygons which lie in the positive real half of the complex plane, and are again symmetric about the real axis. Because of this symmetry we need only consider the positive imaginary half plane in the discussion that follows. Again, we wish to maximize $F(\eta)$ under the constraint that the ellipse E has all vertices of the $\ln$ -transformed polygon either inside or on the boundary. The first A and B are again the coordinates of the vertex with largest imaginary component after the $\ln$ -transformation. The c and η_c above can be computed for this vertex with each other one. Then the other vertex which gives the smallest value of $\eta_c > B$, call it η_{c_1} , determines the next values A_1 and B_1 of A and B which are taken on for $\eta \geq \eta_{c_1}$ but less than some η_{c_2} . Similarly determine the next values η_{c_2} , A_2 and B_2 , which are used for $\eta \geq \eta_{c_2}$ but less than some η_{c_3} . Continuing this process as long as necessary gives the sequence of vertices which the boundary of the deforming ellipse passes through so that all vertices are always in the ellipse or on the boundary. It also gives the sequence of values of η at which the boundary of the ellipse passes from one vertex to the next. This defines the $F(\eta)$ for spectra in convex polygons. It is continuous on the region of interest.

The maximum of $F(\eta)$ on $B < \eta < \eta_{\text{sup}}$ can be found from the derivative

$$\frac{dF}{d\eta} = \frac{\left[\left(\frac{\pi}{2} - \eta \right) \tan \eta - \ln(2 \cos \eta) + A \frac{\left(\frac{\pi}{2} B^2 - \eta^3 \right)}{(\eta^2 - B^2)^{3/2}} \right]}{2[\ln(2 \cos \eta) + A\eta/(\eta^2 - B^2)^{1/2}]^2} . \quad (1.8)$$

Potential maxima of F occur at the endpoints of the interval and at points where $\frac{dF}{d\eta}$ passes through zero or, since $\frac{dF}{d\eta}$ has a finite number of jump discontinuities, points where $\frac{dF}{d\eta}$ changes sign. The derivative of F approaches infinity at the left end of the interval and F approaches a finite number at the right end. Numerical trials indicate that there is generally one sign change of the derivative in the interval and it indicates the maximum of F . For spectra with large complex components compared to the real components there may be no sign changes, in which case this method is not applicable.

2. Tests and Results

Tests of the ADI method using parameters generated by the theory presented here for complex spectra were run on the Lyapunov matrix equation

$$A^T X + XA = W \quad (2.1)$$

where A and W are known 20×20 matrices. The main ADI programs were written by Allen Crider, and all programs that were used are listed in Appendix B. The program used

to calculate k' for spectra of A in a general convex polygon was an insignificant variation on the method of the previous section in that the switching values of η were determined during the iteration by checking that all eigenvalues defining the polygon lay inside the ellipse for each value of η and changing the vertex which determined the numbers A and B when necessary to make this so.

The spectra of the matrices A lay in three different types of smallest convex polygon: four-sided or rectangular; six-sided or trapezoidal, specifically, two trapezoids symmetrically placed around the real axis; and ten-sided, which was meant to approximate a rectangle with rounded corners. All of the polygons were symmetric about the real axis and were actually very close to rectangular. Table 1 lists the defining points of the smallest convex polygon for each matrix A tested. The eigenvalues in the rectangular polygons lay evenly spaced along the top and bottom edges of the rectangle, in the trapezoidal polygons they lay along the top and bottom edges of the region and on the real axis, and in the ten-sided polygons they lay on two ellipses, one at each end of the region.

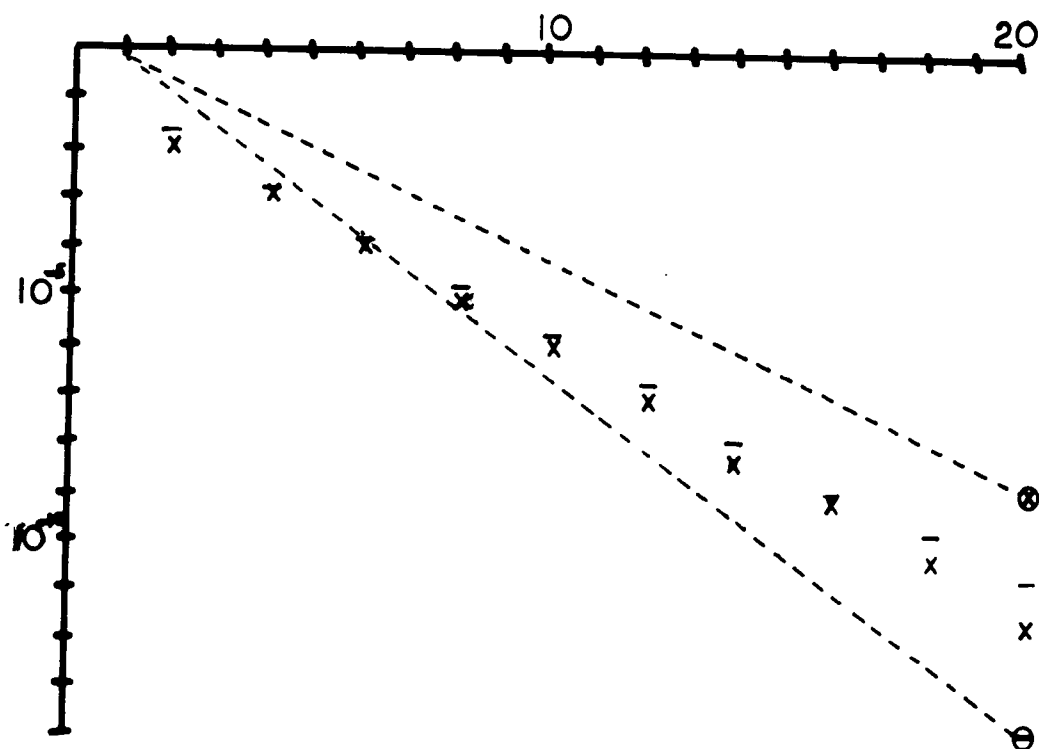
The numerical tests on rectangles were run for complex components increasing until no reasonable η could be found to maximize F , except for the two with smallest real extent for which complex components could be increased until

Table 1. Eigenvalues defining the smallest convex polygon containing the spectrum for each matrix A for which tests of the ADI method were run.

<u>Rectangles</u>	
a) $.2 \pm .05i$, $5 \pm .05i$	r) $.005 \pm .00035i$, $200 \pm .00035i$
b) $.2 \pm .1i$, $5 \pm .1i$	s) $.005 \pm .0007i$, $200 \pm .0007i$
c) $.2 \pm .15i$, $5 \pm .15i$	
d) $.2 \pm .2i$, $5 \pm .2i$	t) $.001 \pm .266 \times 10^{-4}i$, $1000 \pm .266 \times 10^{-4}i$
e) $.2 \pm .25i$, $5 \pm .25i$	u) $.001 \pm .4 \times 10^{-4}i$, $1000 \pm .4 \times 10^{-4}i$
f) $.2 \pm .3i$, $5 \pm .3i$	v) $.001 \pm .533 \times 10^{-4}i$, $1000 \pm .533 \times 10^{-4}i$
	w) $.001 \pm .8 \times 10^{-4}i$, $1000 \pm .8 \times 10^{-4}i$
g) $.1 \pm .025i$, $10 \pm .025i$	x) $.0005 \pm .2 \times 10^{-4}i$, $2000 \pm .2 \times 10^{-4}i$
h) $.1 \pm .05i$, $10 \pm .05i$	y) $.0005 \pm .3 \times 10^{-4}i$, $2000 \pm .3 \times 10^{-4}i$
i) $.1 \pm .075i$, $10 \pm .075i$	
j) $.1 \pm .1i$, $10 \pm .1i$	z) $.0001 \pm .2 \times 10^{-5}i$, $10000 \pm .2 \times 10^{-5}i$
k) $.1 \pm .125i$, $10 \pm .125i$	zz) $.0001 \pm .4 \times 10^{-5}i$, $10000 \pm .4 \times 10^{-5}i$
l) $.1 \pm .15i$, $10 \pm .15i$	
m) $.05 \pm .00917i$, $20 \pm .00917i$	
n) $.05 \pm .01833i$, $20 \pm .01833i$	
o) $.05 \pm .0275i$, $20 \pm .0275i$	
p) $.01 \pm .00095i$, $100 \pm .00095i$	
q) $.01 \pm .0019i$, $100 \pm .0019i$	
<u>Six-sided polygons</u>	
a) $.0149+0i$, $.0179 \pm .003i$, $3.4821 \pm .003i$, $3.4851+0i$	
b) $.0179+0i$, $.0279 \pm .01i$, $3.4921 \pm .01i$, $3.5021+0i$	
c) $.0479+0i$, $.0779 \pm .03i$, $3.5421 \pm .03i$, $3.5721+0i$	
d) $.0579+0i$, $.0679 \pm .01i$, $3.5321 \pm .01i$, $3.5421+0i$	
<u>Ten-sided polygons</u>	
a) $.0149+0i$, $.0155 \pm .0019i$, $.0171 \pm .003i$, $2.0190 \pm .003i$, $2.0206 \pm .0019i$, $2.0212+0i$	
b) $.0179+0i$, $.0199 \pm .0062i$, $.0252 \pm .01i$, $2.0317 \pm .01i$, $2.0369 \pm .0062i$, $2.0390+0i$	
c) $.0479+0i$, $.0539 \pm .0186i$, $.0697 \pm .03i$, $2.0892 \pm .03i$, $2.1050 \pm .0186i$, $2.1111+0i$	
d) $.0579+0i$, $.0599 \pm .0062i$, $.0652 \pm .01i$, $2.0717 \pm .01i$, $2.0769 \pm .0062i$, $2.0790+0i$	

convergence worse than the theoretical bounds assuming complex spectra became apparent. Typical graphs for rectangular spectra of the actual and theoretical error versus number of iterations for optimum parameters computed using real spectral range only and for optimum parameters computed taking into account the imaginary components of the spectrum are shown in Figure IV-3. The important features of these graphs appeared in all the tests on rectangles. In all cases where the imaginary magnitude was small compared to the smallest real component, specifically, for all rectangles tested except those with minimum real component .2 or .1 and imaginary part .15 or .125 or larger, respectively, the actual convergence of the method using parameters for complex spectra was very similar to that using parameters for real spectra for each rectangle (see Figure IV-3, a-d). In these cases, the iterations using parameters for complex spectra had generally less error at each point measured than those using parameters for real spectra. Occasionally, this would not hold for a few iterations, but the differences then were always very small.

The very dramatic difference between the two methods for this data set occurred in the theoretical error bounds. The theoretical error bounds for the parameters for real spectra were consistently smaller than the actual error

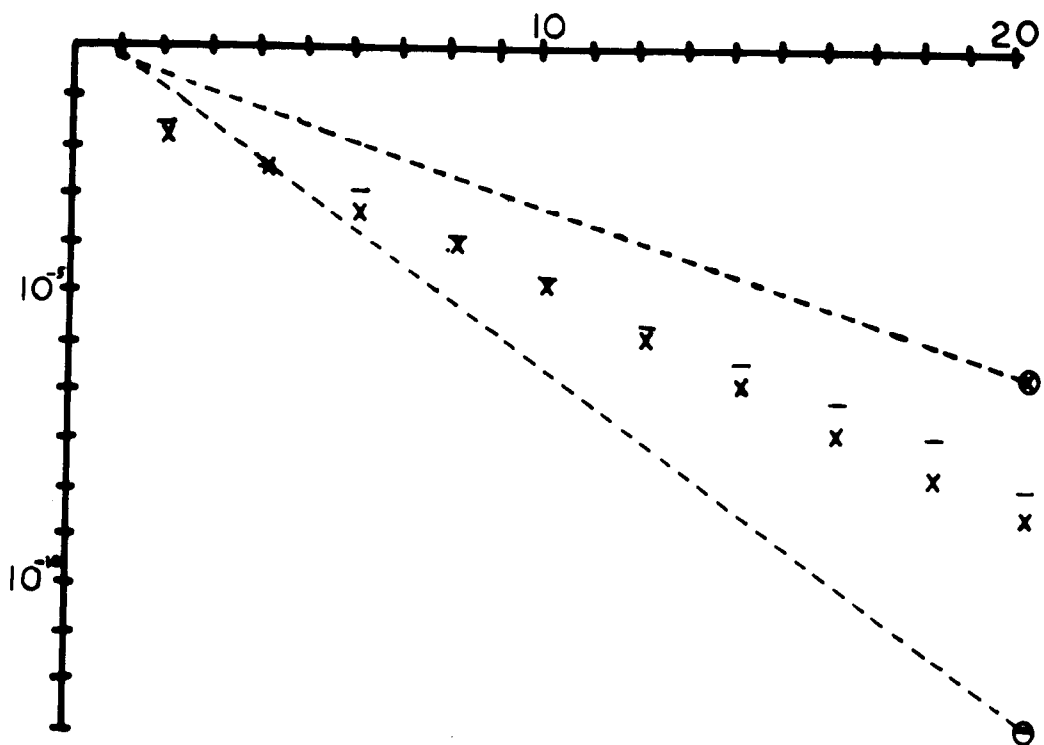


a) maximum absolute complex value .025

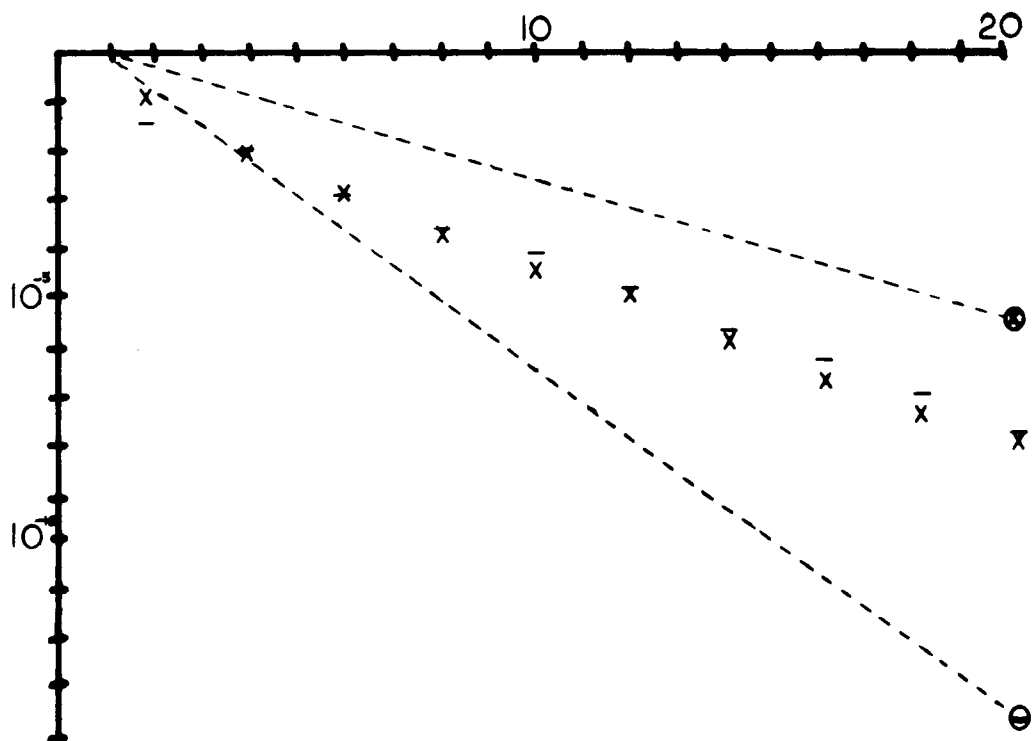
Figure IV-3. Actual error and theoretical error bounds (equation (3.8) of Chapter III, p. 30) for spectra in six rectangles with real range .1 to 10 and maximum absolute complex values as indicated. See text for discussion.

Key for Figures IV-3, IV-4, and IV-5

- -actual error for parameters for real spectra
- ⊕ -theoretical error for parameters for real spectra
- x -actual error for parameters for complex spectra with rectangular bounds
- ⊗ -theoretical error for parameters for complex spectra with rectangular bounds
- 0 -actual error for parameters for complex spectra with polygonal bounds
- -theoretical error for parameters for complex spectra with polygonal bound

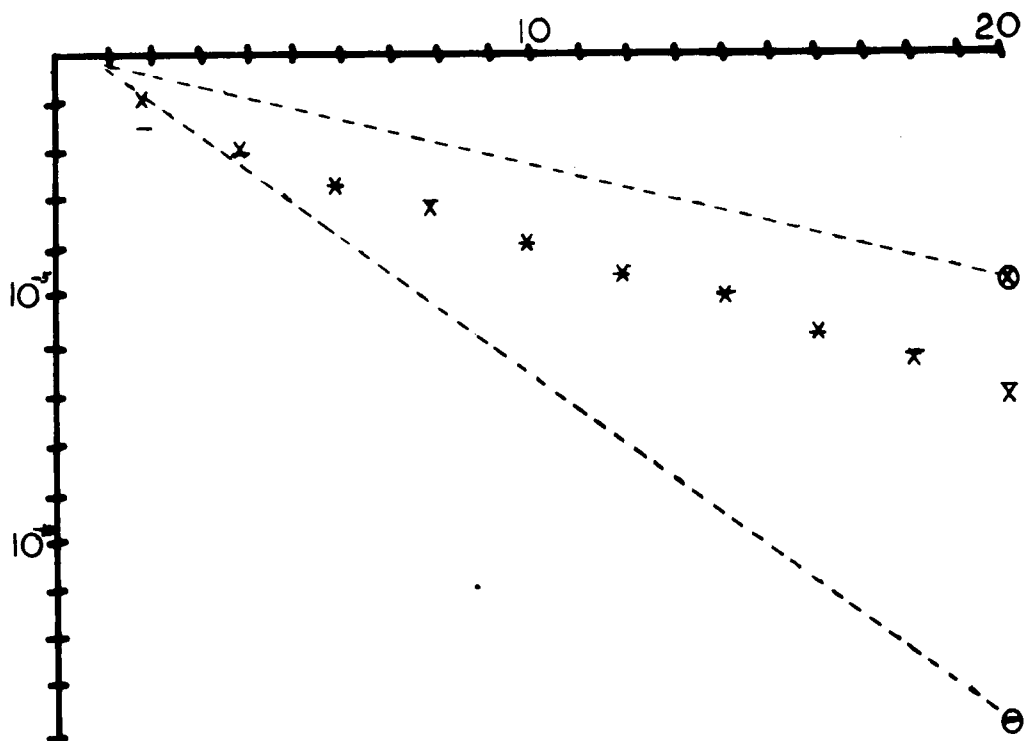


b) maximum absolute complex value .05

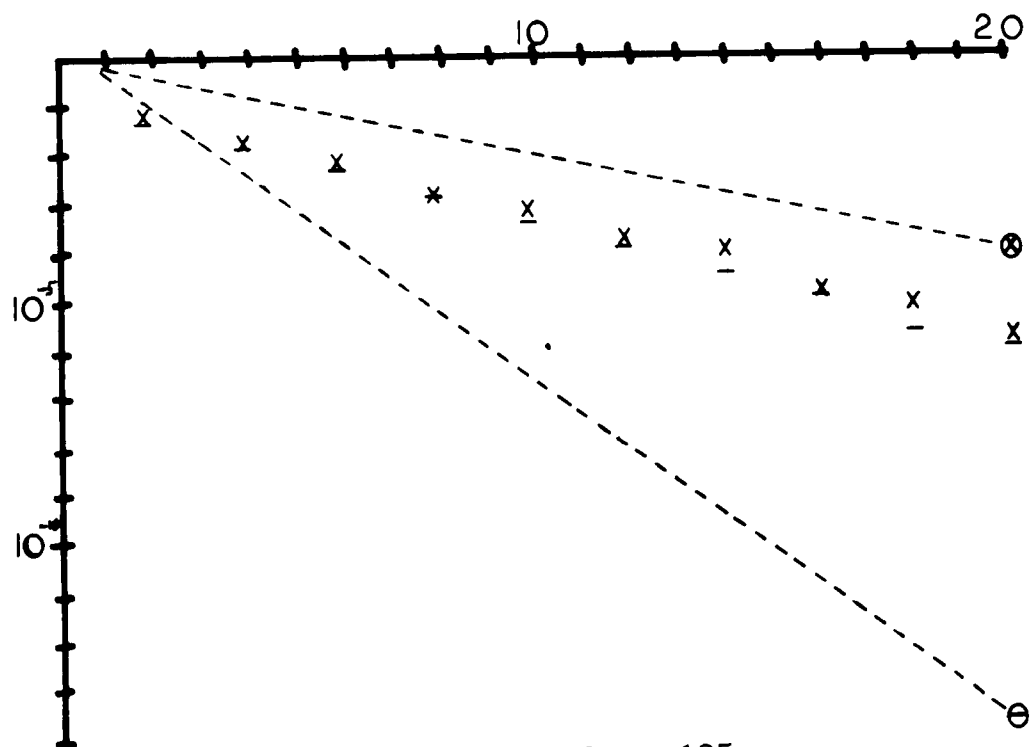


c) maximum absolute complex value .075

Figure IV-3 (Continued)

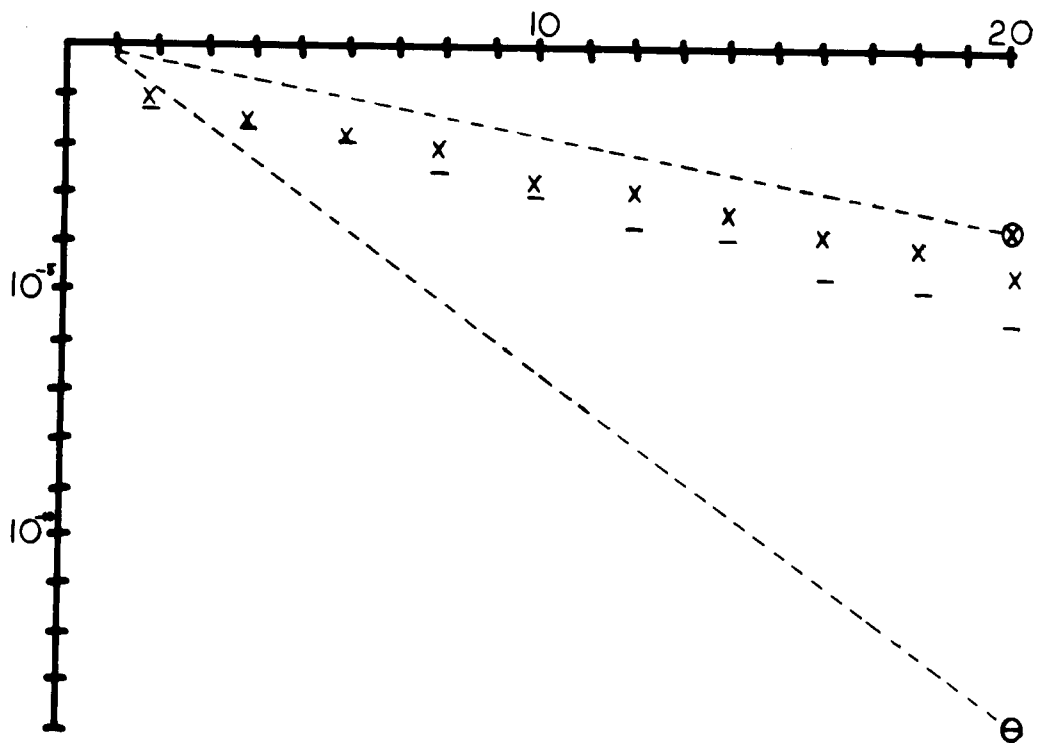


d) maximum absolute complex value .10



e) maximum absolute complex value .125

Figure IV-3 (Continued)



f) maximum absolute complex value .15

Figure IV-3 (Continued)

using either set of parameter values for all but the cases with very small ($\leq .00035$) complex spectral component. The theoretical error bounds for parameters for complex spectra were always good. This result indicated that the theory is quite accurate in these cases.

For the exceptional cases above, where the magnitude of the complex component was almost as large or larger than the magnitude of the smallest real component, the results of the iterations using parameters for complex spectra were worse than those using parameters for real spectra (see Figure IV-3, e and f). Also, the actual error did not necessarily stay within the theoretical bounds for either parameter iteration. The problem occurred when a value of k' was used that was larger than the optimum k' assuming real spectra, k'_r , which had the effect of spreading out the iteration parameters ω_j .

Figure IV-4 shows some representative results for trapezoidal and ten-sided spectral bounds. The results were similar to those for the rectangular spectra. Actual convergence was very similar for parameters computed assuming real spectra, complex spectra with rectangular bounds, and complex spectra with polygonal bounds, but theoretical error bounds of the parameters assuming complex spectra were much more appropriate than those of parameters assuming real spectra. However, as in the case of the rectangular spectra, both actual error and theoretical error bounds

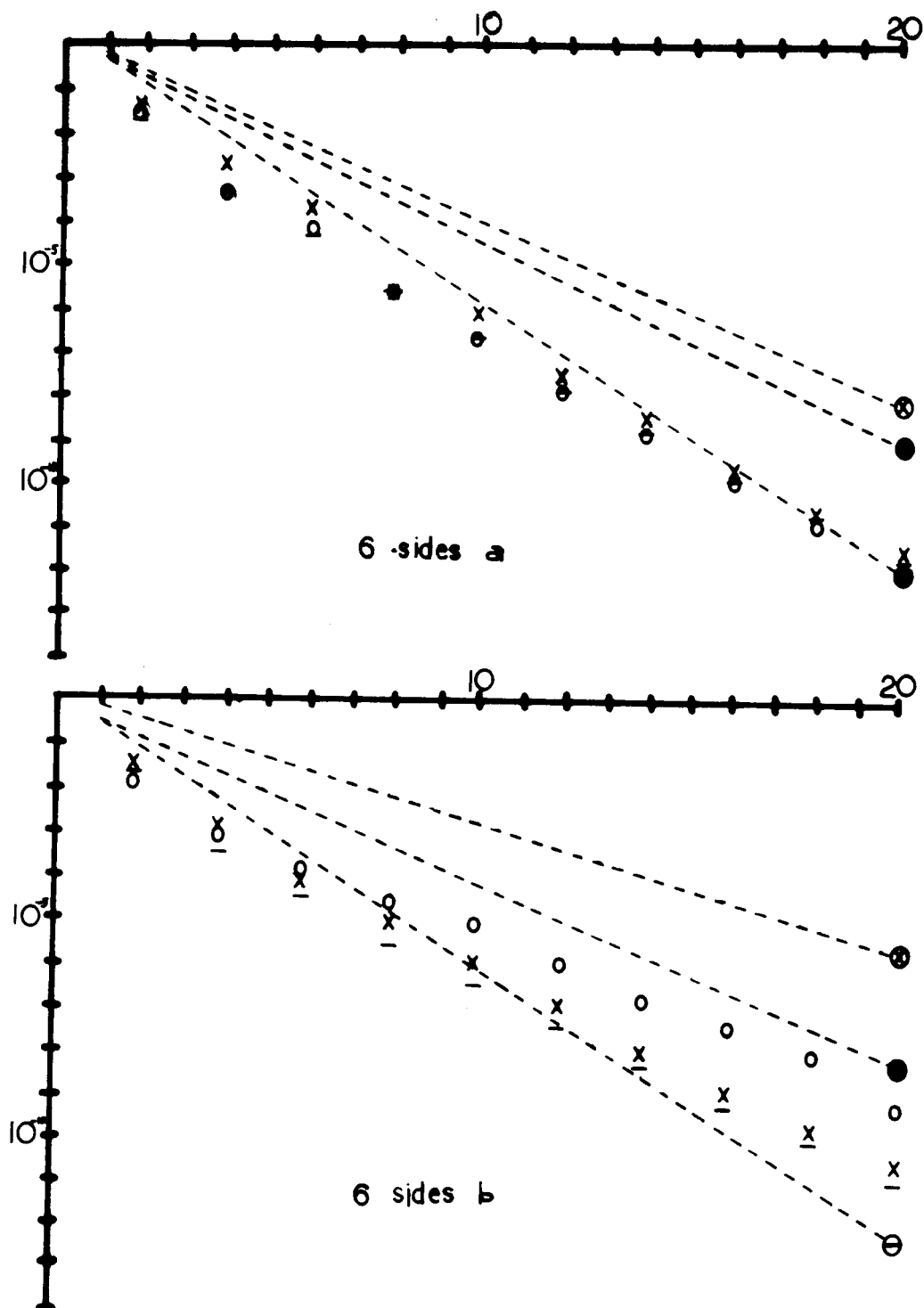


Figure IV-4. Actual error and theoretical error bounds for spectra in the two six-sided regions a and b in Table 1 and the two ten-sided regions a and b in Table 1.* See text for discussion.

*See Figure IV-3, p. 43 for symbol key.

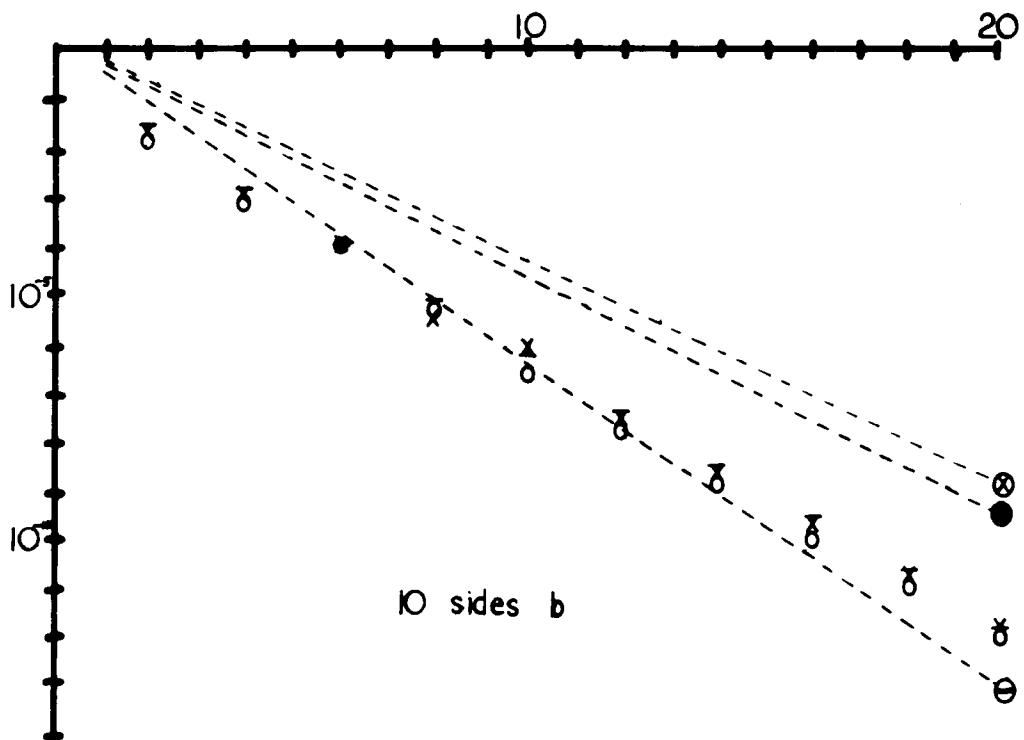
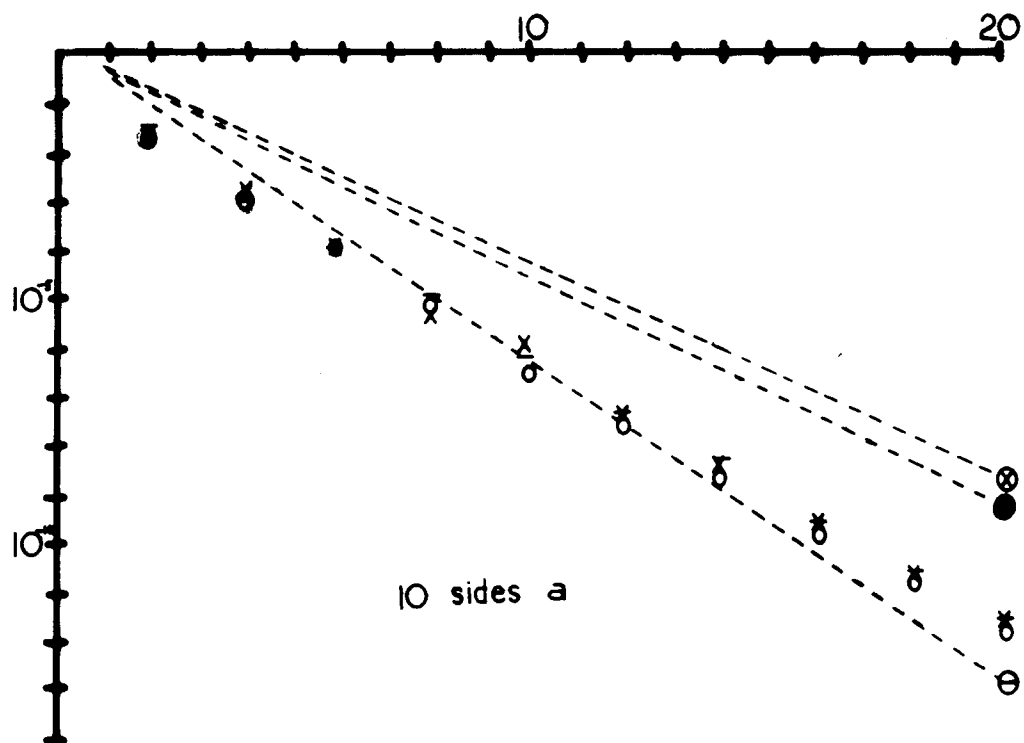


Figure IV-4 (Continued)

were poor for the spectra tested which had larger complex components.

Two other tests were run. One was to compare actual and theoretical convergence rates for the ten-sided spectral regions for a series of k' in an interval around the theoretically optimum k' values. Representative results are shown in Figure IV-5. These tests showed good agreement between the theoretical error bounds for complex spectra and the actual convergence, indicating again the appropriateness of the theory developed here to certain spectral ranges. The discussion of the other test is more appropriate to the next section.

3. Relation to the Theory for Eigenvalues on a Circle

The other test was the investigation of convergence for different sets of iteration parameters ω_j for spectra located on a circle centered on the real axis with ln-symmetric real part. Theory previously developed by Wachspress [personal communication] indicates that the true optimum parameters ω_j should be unity, repeated. Then using equation (2.1) from Chapter 3, p. 20, the error reduction for t double sweeps of ADI iteration should be

$$Q_t^2(z, \omega) = \left| \frac{1-z}{1+z} \right|^{2t} = \left| \frac{1-a}{1+a} \right|^{2t} \quad (3.1)$$

where z lies on the eigenvalue circle which has minimum real value $z = a$ and maximum real value $z = 1/a$, that is

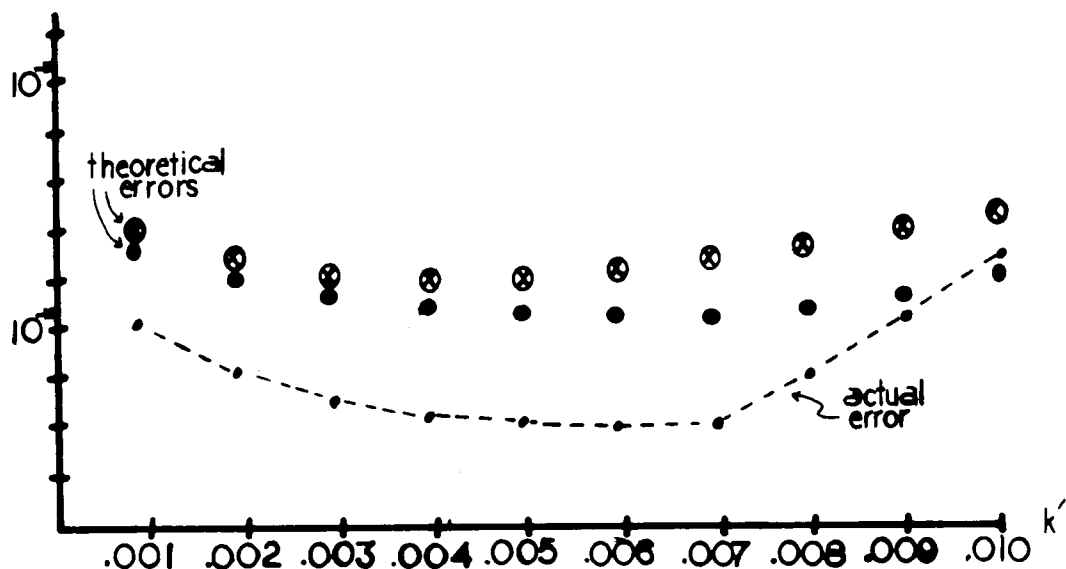


Figure IV-5. Actual error and theoretical error bounds after 20 iterations for spectra in the ten-sided polygonal region a in Table 1 for varying k' . This shows good approximation of the error bounds to the actual error when the spectrum is bounded by a rectangle, and even better approximation when the spectrum is bounded by the ten-sided polygon.*

*See Figure IV-3, p. 43 for symbol key.

$$z = \frac{a + \frac{1}{a}}{2} + \frac{\frac{1}{a} - a}{2} e^{i\theta}. \quad (3.2)$$

The proof, due to Wachspress, that

$$\left| \frac{1-z}{1+z} \right| = \left| \frac{1-a}{1+a} \right| \quad (3.3)$$

is as follows. From (3.2) one obtains

$$\left| \frac{1-z}{1+z} \right| = \left| \frac{2 - (a + \frac{1}{a}) - (\frac{1}{a} - a)e^{i\theta}}{2 + (a + \frac{1}{a}) + (\frac{1}{a} - a)e^{i\theta}} \right|. \quad (3.4)$$

The identities

$$\begin{aligned} 2 - (a + \frac{1}{a}) &= -(\sqrt{a} - \frac{1}{\sqrt{a}})^2 \\ 2 + (a + \frac{1}{a}) &= (\sqrt{a} + \frac{1}{\sqrt{a}})^2 \\ (\frac{1}{a} - a) &= (\frac{1}{\sqrt{a}} - \sqrt{a})(\frac{1}{\sqrt{a}} + \sqrt{a}) \end{aligned} \quad (3.5)$$

then give

$$\left| \frac{1-z}{1+z} \right| = \left| \frac{\sqrt{a} - \frac{1}{\sqrt{a}}}{\sqrt{a} + \frac{1}{\sqrt{a}}} \right| \cdot \left| \frac{(\frac{1}{\sqrt{a}} - \sqrt{a}) + (\frac{1}{\sqrt{a}} + \sqrt{a})e^{i\theta}}{(\frac{1}{\sqrt{a}} + \sqrt{a}) + (\frac{1}{\sqrt{a}} - \sqrt{a})e^{i\theta}} \right|. \quad (3.6)$$

Let

$$\xi = \left| \frac{\frac{1}{\sqrt{a}} - \sqrt{a}}{\frac{1}{\sqrt{a}} + \sqrt{a}} \right| = \left| \frac{1-a}{1+a} \right|. \quad (3.7)$$

Then

$$\left| \frac{1-z}{1+z} \right| = \xi \left| \frac{\xi + e^{i\theta}}{1 + \xi e^{i\theta}} \right| = \xi. \quad (3.8)$$

Only one example of ADI iteration for eigenvalues on a circle was run. This example had a radius of 49.995, $a = .01$, and $1/a = 100$. Theoretical error for the

Wachspress parameters for twenty iterations is

$$\left| \frac{1 - .01}{1 + .01} \right|^{40} \left| \frac{.99}{1.01} \right|^{40} = .45. \quad (3.9)$$

Actual error was slightly less than the theoretical error for the Wachspress parameters for all numbers of iterations tested. Some other sets of parameters gave insignificantly faster convergence, probably due to uneven distribution of the eigenvalues on the circle.

Wachspress' theory of true optimum iteration parameters for eigenvalues on a circle suggest that for spectra with increasing complex components for the same real range, true optimum iteration parameters converge to one, as illustrated in Figure IV-6. The results of the computations in this thesis support that theory. Table 2 lists the optimum iteration parameters ω_j for eight iterations for the series of rectangles for which convergence rates were shown in Figure IV-3. For spectra with smaller magnitudes of the complex components, the optimum parameters start to converge towards one as the complex spectral components increase. However, when the complex components get large enough so that the theory developed here appears to break down, the optimum parameters begin to diverge and the convergence rate of the iteration deteriorates. Table 3 illustrates convergence of optimum iteration parameters for a series of rectangles for which the theory is very appropriate.

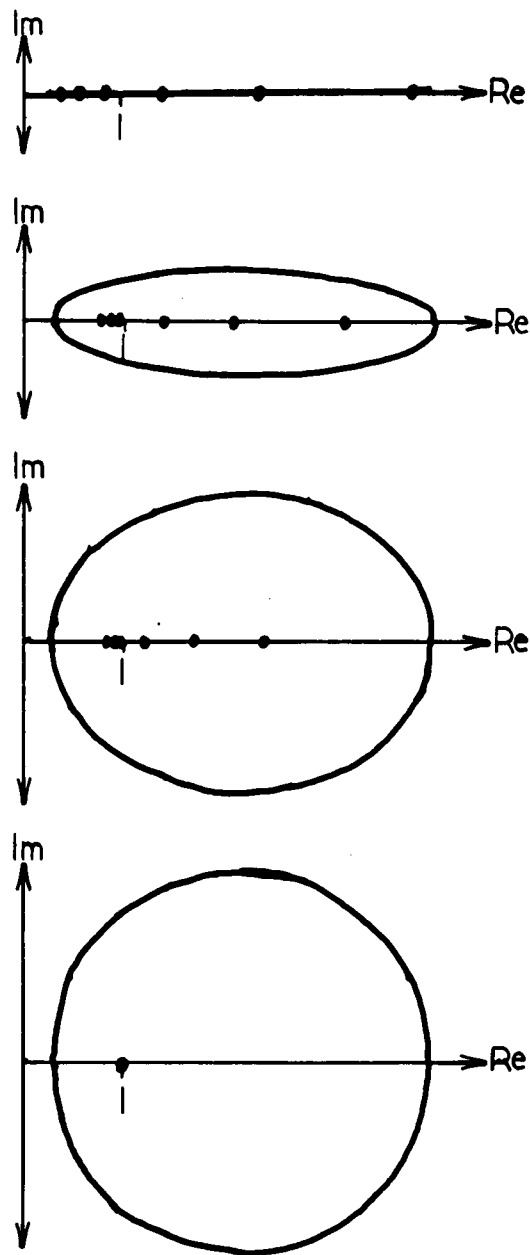


Figure IV-6. Qualitative illustration of expected behavior of true optimum iteration parameters (dots on real axis) for elliptical spectra with increasing complex component. The limiting cases are spectra on a circle and spectra on the real line.

Table 2. Optimum iteration parameters ω_j for eight iterations for the series of rectangles with real range .1 to 10 in Table 1. Convergence rates for these rectangles with these parameters were shown in Figure IV-3, p. 43. See discussion in text.

Max. Magnitude of the Imaginary Component	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
0	9.34	5.88	3.00	1.45	.69	.33	.17	.11
.025	9.21	5.37	2.52	1.12	.50	.22	.10	.06
.05	9.20	5.35	2.50	1.11	.49	.22	.10	.06
.075	9.25	5.52	2.67	1.21	.56	.25	.12	.07
.10	9.32	5.82	2.95	1.41	.66	.32	.16	.10
.125	9.41	6.23	3.36	1.71	.86	.43	.23	.16
.15	9.52	6.73	3.94	2.16	1.16	.64	.37	.26

Table 3. Optimum iteration parameters ω_j for eight iterations for the series of rectangles with real range .05 to 20 in Table 1. See discussion in text.

Max. Magnitude of the Imaginary Component	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6	ω_7	ω_8
0	18.05	9.44	3.95	1.58	.63	.25	.11	.06
.00917	17.69	8.37	3.17	1.16	.42	.16	.06	.03
.01833	17.54	8.00	2.91	1.02	.36	.13	.05	.02
.0275	17.48	7.81	2.80	.97	.34	.12	.04	.02

CHAPTER 5

CONCLUSION

The essence of this work is the extension of the existing ADI theory for real spectra to include some complex spectral regions with small complex components. This was accomplished by finding families of complex spectral regions for which the parameters of the existing theory could be seen to be optimal by a slight extension of that theory. These regions are the lemniscates described in Chapter III. Complex spectra that can be fit within regions approximating those in these families are expected to yield to this analysis. Such spectral regions will approximate regions where $1/\lambda$ is in the region whenever λ is. The spectrum must also have positive real part. The numerical tests indicate that it is necessary for the complex components of the spectrum to be quite small in magnitude compared to the real components. The success of this analysis can be seen in part in the somewhat improved convergence rates of ADI iterations when parameters computed by this method are used on matrices with appropriate spectra. This analysis has been most successful in providing a more accurate error bound for ADI iteration for matrices with appropriate spectra.

While this work is significant in being one of the first investigations of the behavior of the ADI method for iteration matrices with complex spectra, much more work should be done on the subject. The theory presented here only applies in cases of small complex components and its applicability could be more clearly delimited. In particular, the approximations made have not been carefully examined to determine when they break down. Unsuitable approximations are probably responsible for the poor performance of the scheme when the complex components of the spectrum are increased.

Further work needs to be done also on the theory for spectra in more general regions in the complex plane. Until now the limiting cases of spectra on a circle and of spectra on the real line had been investigated. This work considers the case of small complex deviations from the real line. It shows that the optimum iteration parameters ω_j begin to converge when small complex components of the spectra are introduced, as expected from the true optimum parameters for the circular spectra. However, behavior of ADI iteration for the whole range of spectra between the real line with small complex perturbations and the circle is currently unexplored. Perhaps investigation into this would yield a more unified and complete theory of ADI iteration.

LIST OF REFERENCES

List of References

- M. Abramowitz and I. Stegun, *Handbook of Mathematical Functions with formulas, graphs and Tables*, U.S. National Bureau of Standards, 1972.
- R.H. Bartels and G.W. Stewart, A solution of the matrix equation $AX+XB = C$, *Comm. ACM* 15, pp. 820-826, 1972.
- F. Bowman, *Elliptic Functions*, John Wiley & Sons, Inc., 1953.
- N.S. Ellner and E.L. Wachspress, New ADI Model Problem Applications, *Proceedings of the Fall Joint Computer Conference* Nov. 2-6, 1986, pp. 528-534, IEEE Computer Society Press, 1986.
- G.H. Golub, S. Nash, and C. Van Loan, A Hessenberg-Shur method for solution of the problem $AX+XB = C$, *IEEE Trans. Automat. Control* AC24, pp. 909-913, 1979.
- L.A. Hageman and D.M. Young, *Applied Iterative Methods*, Academic Press, 1981.
- H. Hancock, *Elliptic Integrals*, John Wiley & Sons, Inc., 1917.
- H. Hurwitz, Infinitesimal scaling - A new procedure for modeling exterior field problems, *IEEE Trans. Magn.* 20, No. 5, pp. 1918-1923, 1984.
- T.A. Manteuffel, The Tschebychev iteration for nonsymmetric linear systems, *Numer. Math.* 28, pp. 307-327, 1977.
- G. Opfer and G. Schober, Richardson's iteration for nonsymmetric matrices, *Linear Algebra and Its Applications* 58, pp. 343-361, 1984.
- D.W. Peaceman and H.H. Rachford, Jr., The numerical solution of parabolic and elliptic differential equations, *J. Soc. Indust. Appl. Math.* 3, pp. 28-41, 1955.
- D. Rothschild and A. Jameson, Comparison of four numerical algorithms for solving the Liapunov matrix equation, *Int. J. Control* 11, No. 2, pp. 181-198, 1970.
- R.A. Smith, Matrix equation $XA+BX = C$, *SIAM J. Appl. Math.* 16, No. 1, pp. 198-201.

L.S. Trefethen, Near-circularity of the error curve in complex Chebyshev approximation, *J. Approx. Theory* 31, pp. 344-367, 1981.

E.L. Wachspress, Optimum ADI iteration parameters for a model problem, *J. Soc. Indust. Appl. Math.* 10, pp. 339-350, 1962.

_____, Extended application of alternating direction implicit iteration model-problem theory, *J. Soc. Indust. Appl. Math.* 11, No. 4, pp. 994-1016, 1963.

_____, *Iterative Solution of Elliptic Systems*, Prentice Hall, 1966.

_____, "Two-Variable ADI Minimax Problem," Ph.D. dissertation, Rensselaer Polytechnic Institute, 1968.

_____, Generalized ADI preconditioning, *CAMWA* 10, No. 6, pp. 457-461, 1984.

D.M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, 1971.

APPENDICES

APPENDIX A

The following is Rouché's Theorem, a well known result of complex analysis (see any text on functions of complex variables, for example; the statement here is that of Sveshnikov and Tikhonov [1971]).

Theorem: Let the functions $f(z)$ and $\phi(z)$ be analytic in a closed domain $\bar{G}$, with the following inequality valid on the boundary Γ of G :

$$|f(z)|_{\Gamma} > |\phi(z)|_{\Gamma}.$$

Then the total number of zeros in G of the function $F(z) = f(z) + \phi(z)$ is equal to the total number of zeros of the function $f(z)$.

In Theorem 1, p. 16, we use $f(z) = P_t(z')$ and $\phi(z) = -Q_t(z')$ with $P_t(\mu) = Q_t(\mu)$ for a μ outside the lemniscate. Since $f(z) + \phi(z) = P_t(z') - Q_t(z')$ has only t roots and by this theorem they all lie in the lemniscate, the fact that μ is a root of $P_t(z') - Q_t(z')$ leads to a contradiction.

In Theorem 2, p. 21, we use $f(z) = Q_t(z)$ and $\phi(z) = (-1)^t P_t(z)$. Cancellation of terms gives $f(z) + \phi(z) = Q_t(z) - P_t(z)$ a maximal degree of $2t-1$, and the fact that the numerator is odd leads to a contradiction with the conclusions of Rouché's Theorem.

APPENDIX B

The two programs ADITST and ADIGEN were written by Allen Crider to perform ADI iteration on matrices with real eigenvalues, and modified very slightly to allow computation of iteration parameters for matrices with real eigenvalues as described in this thesis. ADITST reads the matrices A and W in the equation $A^T X + XA = W$, calls ADIGEN to compute X, and calls LINPACK routines to calculate the error in X from the true solution C. The matrix C is tridiagonal with 2 on the diagonal and -1 off the diagonal, and was used in the original computation of W. ADIGEN calculates X and is self-explanatory.

The two programs KS and NSBOVL compute iteration parameters for matrices with complex spectra as described in this thesis. KS is used when the spectrum is bounded by a rectangle, and NSBOVL when the spectrum is bounded by a polygon having up to ten sides. Both find the root, if it exists, of the derivative of the function $F(\eta)$ with respect to η (see Chapter IV, equations 1.2, p.35, and 1.8, p.39). In these programs, η is called w. KS is called from ADIGEN, but NSBOVL must be run separately with the real and imaginary parts of the critical eigenvalues, PR and PI, typed in the body of the program.

ADITST

```

      INTEGER I, J, IER, N, NOIT
      REAL*8 A, SCR, w, X, XMC, XNRMER, ROCI, C, CNRMR
      REAL*8 LMAX, LMIN, LMAXI
      DIMENSION A(50,50), SCR(10000), w(50,50), X(50,50), XMC(50,50)
      DIMENSION C(50,50)
      WRITE(5,1)
1  FORMAT(' TYPE:0=CMPLX,1=RL;XXXX;LMAX,LMIN,LMAXI OR NEGS')
      READ (5,*) ROCI,LMAX,LMIN,LMAXI
      READ (10,*) N
      DO 5 I = 1, N
5  READ (10,*) (A(I,J), J = 1, N)
      DO 20 I = 1, N
20 READ (10,*) (w(I,J), J = 1, N)
      XX=.5
      JX=50
      DO 1000 NOIT=2,24,2
      CALL ADIGEN(X,A,w,N,JX,ROCI,NOIT,XX,SCR,IER,LMAX,LMIN,LMAXI)
      IF (IER .EQ. 0) GO TO 40
      WRITE (22,30) IER
C 30 FORMAT (' ERROR ENCOUNTERED IN ADIGEN; IER = ', I3//)
C 40 DO 45 I = 1, N
C 40 CONTINUE
C 45 WRITE (22,50) (X(I,J), J = 1, N)
50 FORMAT (8F15.8)
      DO 60 I = 1, N
      DO 60 J = 1, N
      C(I,J) = 0.0
60 XMC(I,J) = X(I,J)
      C(1,1) = 2.0
      XMC(1,1) = XMC(1,1) - 2.
      C(1,2) = -1.0
      XMC(1,2) = XMC(1,2) + 1.
      C(N,N) = 2.0
      XMC(N,N) = XMC(N,N) - 2.
      C(N,N-1) = -1.0
      XMC(N,N-1) = XMC(N,N-1) + 1.
      DO 61 I = 2, N-1
      C(I,I-1) = -1.0
      XMC(I,I-1) = X(I,I-1) + 1.
      C(I,I) = -2.0
      XMC(I,I) = X(I,I) - 2.
      C(I,I+1) = -1.0
61 XMC(I,I+1) = X(I,I+1) + 1.
      CALL VNRMF2(XMC,N,50,XNRMER)
C  CALL VNRMS1(XMC,N,XNRMER)
C  WRITE(25,62)XNRMER
      CALL VNRMF2(C,N,50,CNRMER)
      CNRMR=XNRMER/CNRMER
      WRITE(25,62)CNRMER
62 FORMAT(' ACTUAL ERRUR (A-E OVER C-E)= ',E15.8)
1000 CONTINUE
      STOP
      END

```

ADIGEN

```

C      SUBROUTINE ADIGEN(X,A,W,N,IDIM,RI,J,EPS,WK,IER,LMXS,LMNS,LMXIS)
C
C      THIS SUBROUTINE USES AN ADI GENERALIZATION TO OBTAIN AN APPROX-
C      IMATE SOLUTION TO THE MATRIX EQUATION
C
C      (A-TRANPOSE) * X + X * A = W
C
C      FOR GIVEN N BY N MATRICES A AND W. THE REQUIRED INPUTS TO THE
C      SUBROUTINE ARE
C
C      X:    MATRIX USED FOR SOLUTION OF THE EQUATION
C      A:    MATRIX FROM EQUATION
C      W:    MATRIX FROM EQUATION
C      N:    DIMENSION OF THE MATRICES A, W, AND X
C      IDIM: ROW DIMENSION OF THE MATRICES A, W, AND X AS DE-
C            CLARED IN THE CALLING PROGRAM
C      EPS:   SPECIFIED ERROR TOLERANCE
C      WK:    ARRAY USED FOR INTERMEDIATE CALCULATIONS; REQUIRED
C            DIMENSION IS AT LEAST 4*N**2
C      IER:   ERROR CODE USED BY SUBROUTINE TO INDICATE THAT AN
C            ERROR HAS OCCURRED; VALUES ARE
C            J:   NO ERRORS ENCOUNTERED
C            102: CALL TO DGED1 WOULD HAVE RESULTED IN DI-
C                 VISION BY ZERO; NO SOLUTION CALCULATED
C
C      SUBROUTINES REQUIRED BY ADIGEN:
C      DGEFA: LINPACK SUBROUTINE WHICH FACTORS A MATRIX
C      DGED1: LINPACK SUBROUTINE WHICH CALCULATES THE INVERSE OF
C            A MATRIX
C
C      OTHER VARIABLES USED IN THE SUBROUTINE ARE
C
C      ABSUM: SUM OF THE ABSOLUTE VALUES OF THE ENTRIES OF A
C            ROW OF THE MATRIX A OR INV(A)
C      D:     ARRAY REQUIRED BY DGED1, NOT USED BY ADIGEN
C      I, II, IJ, K: USED AS INDICES FOR DO LOOPS
C      INFO:  CODE RETURNED BY DGEFA TO INDICATE WHETHER DIVISION
C            BY ZERO WOULD RESULT FROM CALL TO DGED1
C      RI:    WHEN 0, PROGRAM USES COMPLEX THEORY, WHEN 1, REAL
C      J:     NUMBER OF ITERATIONS TO BE PERFORMED
C      KP:    LMIN/LMAX
C      KP4:   KP/4
C      LMAX:  BOUND ON MAXIMUM EIGENVALUE OF A
C      LMIN:  BOUND ON MINIMUM EIGENVALUE OF A
C      LMXS,LMNS,LMXIS: ALLOW BOUNDS ON EIGENVALUES TO BE PRESET
C            IN CALLING PROGRAM. IF THEY ARE PRESET NEGATIVE,
C            THE COMPUTED VALUE IS USED.
C      PJ:    B * WJ
C      RJ:    (2 * IJ - 1)/(2 * J), WHERE IJ IS THE NUMBER OF THE
C            CURRENT ITERATION
C      RSUM:  SUM OF THE ELEMENTS OF A ROW THE MATRIX A
C      SING:  USED TO INDICATE POSSIBILITY OF A BEING SINGULAR
C      WJ:    INTERMEDIATE VALUE IN CALCULATION OF ITERATION
C            PARAMETER
C
C      INTEGER I, IDIM, IER, II, IJ, INFO, J, K, N, SING
C      REAL*8 A, ABSUM, D, EPS, KP, KP4, LMAX, LMAX1, LMIN, OFSUM, PJ,
C      RJ, RSUM, W, WJ, WK, X, AINM, BINM, SIG, TERR,
C      LMXS,LMNS,LMXIS
C      DIMENSION A(IDIM,1), D(2), W(IDIM,1), WK(N,N,4), X(IDIM,1)
C

```

C CALCULATE BOUND ON THE MAXIMUM EIGENVALUE OF A AND TEST FOR POS-
C SIBLE SINGULARITY.
C

```

DO 10 I = 1, N
  DO 10 IJ = 1, N
    WK(I,IJ,2) = A(I,IJ) - A(IJ,I)
10  WK(I,IJ,1) = A(I,IJ) + A(IJ,I)
    LMAX = 0.
    LMAXI = 0.
    SING = 1
    DO 20 I = 1, N
      ABSUM = 0.
      OFSUM = 0.
      RSUM = 0.
      DO 15 II = 1, N
        ABSUM = ABSUM + ABS(WK(I,II,1))
        OFSUM = OFSUM + ABS(WK(I,II,2))
15  RSUM = RSUM + WK(I,II,1)
        OFSUM = OFSUM - ABS(WK(I,II,2))
        IF (ABS(RSUM) .GT. 0.1D-08) SING = 0
        IF (ABSUM .GT. LMAX) LMAX = ABSUM
        IF (OFSUM .GT. LMAXI) LMAXI = OFSUM
20  CONTINUE
    LMAX = LMAX / 2.
    LMAXI = LMAXI / 2.
    IF (LMXS.GT.0.0) LMAX=LMXS
    IF (LMXIS.GE.0.0) LMAXI=LMXIS

```

C FIND INVERSE OF A TO OBTAIN BOUND ON MINIMUM NON-ZERO EIGENVALUE.
C

```

IF (SING. EQ. 0) GO TO 50
DO 40 I = 1, N
  DO 40 II = 1, N
    WK(II,II,1) = WK(I,II,1) + 1./N
40  WK(II,II,1) = WK(I,II,1) + 1./N
50 CALL DGEFA (WK(1,1,1), N, N, WK(1,1,2), INFO)
IF (INFO. EQ. 0) GO TO 60
IER = I02
RETURN
60 CALL DGED1 (WK(1,1,1), N, N, WK(1,1,2), D, WK(1,1,3), 1)

```

C CALCULATE BOUND ON MINIMUM NON-ZERO EIGENVALUE OF A
C

```

LMIN = 0.
DO 80 I = 1, N
  ABSUM = 0.
  DO 70 II = 1, N
    ABSUM = ABSUM + ABS(WK(I,II,1))
70  IF (ABSUM .GT. LMIN) LMIN = ABSUM
80  CONTINUE
    LMIN = 0.5/LMIN
    IF (LMNS.GT.0.0) LMIN=LMNS
    WRITE (25,82) LMIN,LMAX,LMAXI
82 FORMAT (' LMIN = ',F12.5,' LMAX = ',F12.5,' LMAXI = ',F12.5)

```

C CALCULATE OPTIMUM KP, SIGMA
C

```

KP = LMIN/LMAX
SIG = 0.0
IF (RI. GT. 0.5) GO TO 1110
CALL KS(LMAX,LMIN,LMAXI,KP,SIG)
1110 WRITE(25,1111)KP,SIG,RI
1111 FORMAT(' KP=',F16.13,' SIGMA=',F9.6,' RI = ',F4.1,'(O=C,I=R)')

```

C CALCULATE NUMBER OF ITERATIONS TO BE PERFORMED
C

```

KP4 = KP/4.

```

```

C      AINM=4.0*(0.5-SIG*LOG(KP4)/3.14159)
C      BINM=9.8696044/(2.0*LOG(KP4))+3.14159*SIG
C      J = AINT((LOG(EPS)-LOG(AINM))/BINM+1)
      WRITE (25,85) J
85  FORMAT (' NUMBER OF ITERATIONS TO BE PERFORMED = ',I3)
C
C      PERFORM ITERATIONS
C
      DO 210 IJ = 1, J
C
C      CALCULATE PARAMETERS FOR ITERATION
C
      RJ = (2. * IJ - 1.)/(2. * J)
      WJ = 2. * KP4 ** KJ * (1. + KP4 ** (2. * (1. - RJ))) /
1      (1 + KP4 ** (2. * RJ))
      PJ = LMAX * WJ
      WRITE (25,87) PJ
87  FORMAT (' PJ = ',F10.5)
C
C      CALCULATE HJ = INV(PJ * I + A), STORE IN WK(*,*,1)
C
      DO 90 I = 1, N
        DO 90 II = 1, N
          90      WK(I,II,1) = A(I,II)
      DO 100 I = 1, N
        100      WK(I,1,1) = WK(I,1,1) + PJ
      CALL DGEFA (WK(1,1,1), N, N, WK(1,1,2), INFO)
      IF (INFO .EQ. 0) GO TO 110
      IER = 102
      RETURN
110      CALL DGED1 (WK(1,1,1), N, N, WK(1,1,2), D, WK(1,1,3), 1)
C
C      CALCULATE 2 * PJ * HJ, STORE IN WK(*,*,2)
C
      DO 120 I = 1, N
        DO 120 II = 1, N
          120      WK(I,II,2) = 2. * PJ * WK(I,II,1)
C
C      CALCULATE KJ = 2 * PJ * (HJ-TRANSPOSE) * W * HJ, STORE IN
C      WK(*,*,4)
C
      DO 130 I = 1, N
        DO 130 II = 1, N
          WK(I,II,3) = 0.
          DO 130 K = 1, N
            130      WK(I,II,3) = WK(I,II,3) + W(I,K) * WK(K,II,2)
      DO 140 I = 1, N
        DO 140 II = 1, N
          WK(I,II,4) = 0.
          DO 140 K = 1, N
            140      WK(I,II,4) = WK(I,II,4) + WK(K,I,1) * WK(K,II,3)
C
C      CALCULATE GJ = I - 2 * PJ * HJ, STORE IN WK(*,*,2)
C
      DO 150 I = 1, N
        DO 150 II = 1, N
          150      WK(I,II,2) = -WK(I,II,2)
      DO 160 I = 1, N
        160      WK(I,1,2) = 1. + WK(I,1,2)
      IF (IJ .NE. 1) GO TO 180
C
C      IF FIRST ITERATION, SET STARTING VALUE FOR X
C
      DO 170 I = 1, N
        DO 170 II = 1, N
          170      X(I,II) = WK(I,II,4)
C

```

```

C  CALCULATE NEXT VALUE FOR X
C
180  DO 190 I = 1, N
      DO 190 II = 1, N
        WK(I,II,3) = 0.
        DO 190 K = 1, N
          WK(I,II,3) = WK(I,II,3) + X(I,K) * WK(K,II,2)
190  DO 200 I = 1, N
      DO 200 II = 1, N
        WK(I,II,1) = 0.
        DO 200 K = 1, N
          WK(I,II,1) = WK(I,II,1) + WK(K,I,2) * WK(K,II,3)
200  DO 210 I = 1, N
      DO 210 II = 1, N
        X(I,II) = WK(I,II,4) + WK(I,II,1)
210
C
C  CALCULATE AND WRITE THEORETICAL ERROR
C
      TERR=EXP(J*(9.8696044/LOG(KP4) + 2.0*3.14159*SIGMA))
      TERR=TERR*4.0**((1.0 - (2.0/3.14159)*SIGMA*LOG(KP4)))
      WRITE(25,350)TERR
350  FORMAT(' THEORETICAL ERROR = ',E15.8)
      RETURN
      END

```

KS

```

CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNSU00010
CCCCCCCCCCCCCCCC SUBROUTINE KS CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNSU00020
      SUBROUTINE KS(AMN, AMN, AMX1, KP, SIG)
      REAL*8 AMN, AMX, AMX1, KP, SIG, WSW, RLLC, AILLC, RLRC, AILRC, C, WA, WB NSU00040
      REAL*8 PROD, WP, KP4, A, B, AINTM, BINTM, AINTF, BINTF, AINTD, BINTD, CINTD NSU00050
      REAL*8 DINTD, EINTD, AFTEST, BFTEST, ADF, BDF, T1, T2, TOL1, TOL2, S
      COMMON WSW, RLLC, AILLC, RLRC, AILRC NSU00070
CCCCCCCCCCCCCCCCCCCC INITIALIZATION CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC
      WSW=1.5707963
      RLLC=(LOG(AMN**2.0+AMX1**2.0)-LOG(AMX*AMN))/2.0
      RLLC=ABS(RLLC) NSU00130
      AILLC=ATAN(AMX1/AMN)
      RLRC=(LOG(AMX**2.0+AMX1**2.0)-LOG(AMX*AMN))/2.0
      AILRC=ATAN(AMX1/AMX)
      WRITE(5,*)RLLC,AILLC,RLRC,AILRC
      T1=RLLC*RLLC+AILLC*AILLC NSU00170
      T2=RLRC*RLRC+AILRC*AILRC NSU00180
      IF(T1.GE.T2) GO TO 10 NSU00190
      C=ABS((RLLC*RLRC-RLRC*KLRC)/(AILLC*AILLC-AILRC*AILRC)) NSU00200
      WSW=(RLLC*KLRC+C+AILLC*AILLC)**0.5 NSU00210
      WRITE(5,6) WSW
      6 FORMAT(' WSW=',F9.6)
      10 WA=AILLC + .0001
      WB=WA+0.001 NSU00260
      IF(DFDW(WA).LT.3.0)GO TO 990 NSU00270
CCCCCCCCCCCCCCCCCCCC FIND GROSS INTERVAL OF ROOT OF DFDW CCCCCCCCCC
      DO 101 I=1,350 NSU00280
      IF (WB.GT.1.5707962) WB=1.5707962
      IF (WA.GT.1.5707962) WA=1.5707962
      IF (WA.EQ.WB) WP=1.5707962
      IF (WA.EQ.WB) GO TO 40
      PROD=DFDW(WA)*DFDW(WB) NSU00290
      IF(PROD.LT.3.0)GO TO 132 NSU00300
      KP=EXP(-1.0*AINTM(WB))
      SIG=WB/LOG(4./KP)
      F=(1.5707963-WB)/LOG(4./KP)
      WRITE(5,91)WA,WB,DFDW(WA),DFDW(WB),F,KP,SIG,Q
      IF(KP.GT.1.0)GO TO 40
      WA=WB NSU00310
      WB=WB+0.001 NSU00320
      101 CONTINUE NSU00330
CCCCCCCCCCCCCCCCCCCC APPLY METHOD OF REGULA FALSI TO INTERVAL CCCCCC
      102 CONTINUE NSU00340
      DO 103 I=1,50 NSU00350
      WP=(WB-WA)/2+WA NSU00390
      WRITE(5,91)WP,WA,WB,DFDW(WP),DFDW(WA),DFDW(WB),FTEST(WA),QQ NSU00400
      91 FORMAT(BF10.6) NSU00410
      IF((WB-WA).LT.10.0E-8)GO TO 30 NSU00420
      PROD=DFDW(WA)*DFDW(WP) NSU00430
      IF(PROD.LT.0) WB=WP NSU00440
      IF(PROD.GT.0) WA=WP NSU00450
      103 CONTINUE NSU00460
CCCCCCCCCCCCCCCCCCCC FIND W FOR MAX KP IF THIS OCCURS BEFORE ROOT CCCCCC
      C 104 KP=EXP(-1.0*AINTM(WP))
      C IF(KP.LT.AMN/AMX) GO TO 40
      C WP=WP-0.0001
      C GO TO 104
      C 103 CONTINUE
      C WA=WA+0.0001
      C IF(EXP(-1.0*AINTM(WA)).LE.AMN/AMX) GO TO 103
      C WP=WA-0.0001

```

```

40 WRITE(5,95)WP
CCCCCCCCCCCCCCCCCCCC COMPUTE KP,SIG.  DIAGNOSTICS  CCCCCCCCCCCCCCCCCCCCCC
30 KP=EXP(-1.0*AINTM(WP))
   SIG=WP/LUG(4./KP)
   GO TO 999
990 WRITE(5,92)WA
   GO TO 999
997 WRITE(5,93)WA
   GO TO 999
998 WRITE(5,94)WB
92  FORMAT(' DFDW OF WA LT 0',F9.7)
93  FORMAT(' FTEST OF WA LT 0',F9.7)
94  FORMAT(' FTEST OF WB LT 0',F9.7)
95  FORMAT(' KP IS MAX KP POSSIBLE ',F9.6)
96  FORMAT(' WINT =',E10.5)
999 WINT=WB-WA
   WRITE(5,96) WINT
   RETURN
   END
NSU00640
NSU00650
NSU00660
NSU00670
CCCCCCCCCCCCCCCCCCCC FUNCTION SUBPROGRAMS  CCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCCCCCCCCCCCCC FUNCTION AINTM  CCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCCCCCCCCCCCCC FUNCTION AINTM  CCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCCCCCCCCCCCCC FUNCTION AINTM  CCCCCCCCCCCCCCCCCCCCCC
DOUBLE PRECISION FUNCTION AINTM(W)
IMPLICIT REAL*8 (A-H,O-Z)
COMMON WSW,RLLC,AILLC,RLRC,AILRC
A=RLLC
B=AILLC
IF(W.GE.WSW) A=RLRC
IF(W.GE.WSW) B=AILRC
AINTM=2.0*((A*B)/((W*W-B*B)**0.5)+LOG(COS(W)))
RETURN
END
NSU00680
NSU00690
CCCCCCCCCCCCCCCCCCCC FUNCTION FTEST  CCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCCCCCCCCCCCCC FUNCTION FTEST  CCCCCCCCCCCCCCCCCCCCCC
DOUBLE PRECISION FUNCTION FTEST(W)
IMPLICIT REAL*8 (A-H,O-Z)
COMMON WSW,RLLC,AILLC,RLRC,AILRC
AFTEST=RLLC
BFTEST=AILLC
IF(W.GT.WSW) AFTEST=RLRC
IF(W.GT.WSW) BFTEST=AILRC
AI=(W*W-BFTEST*BFTEST)
AINTF=(W*W-BFTEST*BFTEST)**0.5
BINTF=LOG(COS(W))
FTEST=AINTF*BINTF+AFTEST*W
RETURN
END
NSU00710
NSU00720
NSU00730
NSU00740
NSU00750
NSU00760
NSU00770
NSU00790
NSU00800
NSU00810
NSU00820
NSU00830
CCCCCCCCCCCCCCCCCCCC FUNCTION DFDW  CCCCCCCCCCCCCCCCCCCCCC
CCCCCCCCCCCCCCCCCCCC FUNCTION DFDW  CCCCCCCCCCCCCCCCCCCCCC
DOUBLE PRECISION FUNCTION DFDW(W)
IMPLICIT REAL*8 (A-H,O-Z)
COMMON WSW,RLLC,AILLC,RLRC,AILRC
ADF=RLLC
BDF=AILLC
IF(W.GT.WSW) ADF=RLRC
IF(W.GT.WSW) BDF=AILRC
AINTD=(W*W-BDF*BDF)**0.5
BINTD=LOG(2.0*COS(W))
CINTD=((ADF*(1.5707963*(BDF*BDF)-W*W*W))/(AINTD*AINTD*AINTD))
DINTD=(1.5707963-W)*TAN(W)-BINTD+CINTD
EINTD=BINTD+ADF*W/AINTD
DFDW=DINTD/(2.0*EINTD*EINTD)
RETURN
END
NSU00850
NSU00860
NSU00870
NSU00880
NSU00890
NSU00900
NSU00920
NSU00930
NSU00940
NSU00950
NSU00960
NSU00970

```

NSBOVL

FILE: NSBOVL FORTRAN A

VM/SP CONVERSATIONAL MONITOR SYSTEM

```

.CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNSB00010
REAL*8 KP,SIGMA,C,WA,WBNSB00020
REAL*8 PROD,WP,KP4,A,B,AINTM,BINTM,AINTF,BINTF,AINTD,BINTD,CINTD NSB00030
REAL*8 DINTD, EINTD,AFTEST,BFTEST,ADF,BDF,S NSB00040
REAL*8 PN(6) NSB00050
REAL*8 PR(6),PI(6) NSB00060
COMMON PR,PI NSB00070
CCCCCCCCCCCCCCCCCCCC INITIALIZATION CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCNSB00080
WHERE GOES INITIALIZATION OF PR(I),PI(I). WA=MAX(PI(I))+.0001 NSB00090
PR(1)=2.367687 NSB00100
PI(1)=0.0000 NSB00110
PR(2)=2.215158 NSB00120
PI(2)=0.301289 NSB00130
PR(3)=1.953288 NSB00140
PI(3)=0.378504 NSB00150
PR(4)=2.364126 NSB00160
PI(4)=0.004927 NSB00170
PR(5)=2.366705 NSB00180
PI(5)=0.003037 NSB00190
PR(6)=2.367687 NSB00200
PI(6)=0.0000 NSB00210
WRITE(5,1) NSB00220
1 FORMAT(' ENTER WA') NSB00230
READ(5,*) WA NSB00240
WB=WA+.001 NSB00250
IF(DFDW(WA).LT.0.0)GO TO 990 NSB00260
CCCCCCCCCCCCCCCCCCCC FIND GROSS INTERVAL OF ROOT OF DFDW CCCCCCCCC NSB00270
DO 101 I=1,350 NSB00280
IF (WB.GT.1.5707962) WB=1.5707962 NSB00290
IF (WA.GT.1.5707962) WA=1.5707962 NSB00300
IF (WA.EQ.WB) WP=1.5707962 NSB00310
IF (WA.EQ.WB) GO TO 40 NSB00320
PROD=DFDW(WA)*DFDW(WB) NSB00330
IF(PROD.LT.0.0)GO TO 102 NSB00340
WA=WB NSB00350
WB=WB+.001 NSB00360
101 CONTINUE NSB00370
CCCCCCCCCCCCCCCCCCCC APPLY METHOD OF REGULA FALSI TO INTERVAL CCCCC NSB00380
102 CONTINUE NSB00390
DO 100 I=1,50 NSB00400
WP=(WB-WA)/2+WA NSB00410
IF((WB-WA).LT.10.0E-8)GO TO 30 NSB00420
PROD=DFDW(WA)*DFDW(WP) NSB00430
IF(PROD.LT.0) WB=WP NSB00440
IF(PROD.GT.0) WA=WP NSB00450
100 CONTINUE NSB00460
40 WRITE(5,95)WP NSB00470
CCCCCCCCCCCCCCCCCCCC COMPUTE KP,SIG. DIAGNOSTICS CCCCCCCCCCCCCCCCC NSB00480
30 GO TO 999 NSB00490
990 WRITE(5,92)WA NSB00500
GO TO 999 NSB00510
92 FORMAT(' DFDW OF WA LT 0',F9.7) NSB00520
93 FORMAT(' FTEST OF WA LT 0',F9.7) NSB00530
94 FORMAT(' FTEST OF WB LT 0',F9.7) NSB00540
95 FORMAT(' KP IS MAX KP POSSIBLE ',F9.6) NSB00550

```

VM/SP CONVERSATIONAL MONITOR SYSTEM

72

VITA

Nancy G. Saltzman was born on June 23, 1958, and was raised in Ithaca, N.Y. She received a Bachelor of Arts degree from Cornell University in 1980 with majors in physics and geology and a minor in applied mathematics. After working as a geophysicist for Amoco Production Company, she returned to school in 1983 to study for a Master's of Science degree in mathematics at The University of Tennessee, Knoxville. She was awarded this degree in 1987 after travelling extensively in Europe with her husband while he was on a fellowship in 1984 and 1985.