WAPD-TM-537

AEC RESEARCH AND
DEVELOPMENT REPORT

THE CHEBYSHEV POLYNOMIAL METHOD OF ITERATION

JANUARY 1967

CONTRACT AT-11-1-GEN-14

BETTIS ATOMIC POWER LABORATORY, PITTSBURGH, PA.,
OPERATED FOR THE U. S. ATOMIC ENERGY COMMISSION
BY WESTINGHOUSE ELECTRIC CORPORATION

DISCLAIMER

WAPD-TM-537

UC-32: Mathematics and Computers
Special Distribution

CFSTI PRICES

THE CHEBYSHEV POLYNOMIAL METHOD OF ITERATION

H.C. $ 3.00 ; MN .65

L. A. Hageman

January 1967

RELEASED FOR ANNOUNCEMENT
IN NUCLEAR SCIENCE ABSTRACTS

CONTRACT AT-11-1-GEN-14

For sale by the Clearinghouse for Federal Scientific and Technical Information, National Bureau of Standards, U.S. Department of Commerce, Springfield, Virginia.

NOTE

This document is an interim memorandum prepared primarily for internal reference and does not represent a final expression of the opinion of Westinghouse. When this memorandum is distributed externally, it is with the express understanding that Westinghouse makes no representation as to completeness, accuracy, or usability of information contained therein.

BETTIS ATOMIC POWER LABORATORY PITTSBURGH, PENNSYLVANIA

Operated for the U.S. Atomic Energy Commission by
Westinghouse Electric Corporation

TABLE OF CONTENTS

In this report the practical use of the Chebyshev polynomial method of iteration is discussed. The convergence behavior of the Chebyshev method is given and a numerical strategy is described which can be used to estimate the required acceleration parameters. Numerical examples are given and discussed.

THE CHEBYSHEV POLYNOMIAL METHOD OF ITERATION

L. A. Hageman

I. INTRODUCTION

If the eigenvalues $\{\sigma_i\}_{i=1}^{i=n}$ of a real nxn matrix G are ordered such that $|\sigma_n| \leq |\sigma_{n-1}| \leq \cdots \leq |\sigma_2| \leq |\sigma_1|$, then σ_1 is called the dominant eigenvalue of G if $|\sigma_1| > |\sigma_2|$. Many practical problems in applied mathematics require knowledge of this dominant eigenvalue and its associated eigenvector.

A standard iterative method for finding the dominant eigenvalue and its associated eigenvector is the well-known power method. For any matrix G with a dominant eigenvalue, the power method is a convergent process provided, of course, that the initial guess vector has a nonzero component of the dominant eigenvector. However, when the dominance ratio $\bar{\sigma} \equiv |\sigma_2|/|\sigma_1|$, of the matrix G is close to unity, the rate of convergence of the power method is very slow. Thus, one would like to find ways to accelerate the convergence rate of the basic power method.

One such acceleration scheme is the Chebyshev polynomial extrapolation method. The improvement achieved by Chebyshev polynomial extrapolation depends strongly on the properties of the eigenvalues and eigenvectors of the matrix G. Normally, in applying Chebyshev polynomials, it is assumed that the eigenvectors of G span the associated vector space $V_n(C)$ of G and that the eigenvalues of G are real. Often, however, Chebyshev extrapolation improves the rate of convergence even though the eigenvalues are not real and/or the eigenvectors do not span the vector space. For this case, though, the acceleration achieved may be small.

The convergence rate of the power method is uniquely determined by the properties of the matrix G and the initial guess vector; whereas, the convergence rate of the Chebyshev extrapolation method also depends on the choice of three parameters. The optimum parameter values, i.e., those values for the parameters which maximize the rate of convergence, are functions of the domain in the complex plane which contains the eigenvalues of G. Generally, the eigenvalue domain, and hence also the optimum parameters, is not known *a priori*. Thus, estimating the optimum parameter values is an important but often neglected problem in the practical application of the Chebyshev extrapolation method.

The purpose of this report is to discuss the practical use of the Chebyshev polynomial method of iteration. First, we define the method and give the well-known convergence properties of the Chebyshev iteration method assuming that the eigenvectors span the associated vector space and that the eigenvalues are real. We then discuss the convergence behavior of the Chebyshev method when these assumptions on the eigenvalues and eigenvectors are relaxed. Practical numerical means by which to estimate the needed parameters are described and a numerical strategy given. Finally, numerical examples are given and discussed.

Although this report is concerned primarily with the solution of the homogeneous eigenvalue problem, much of what is said is valid also for the inhomogeneous problem. The use of the Chebyshev polynomial method of iteration in the solution of the inhomogeneous matrix problem is discussed briefly in Appendix B.

II. THE CHEBYSHEV POLYNOMIAL METHOD

1. Introduction

Let G be a real nxn matrix with eigenvalues $\left\{\sigma_i\right\}_{i=1}^{i=n}$ and eigenvectors $\left\{\underline{x}_i\right\}_{i=1}^{i=n}$. We assume that the matrix G has a dominant eigenvalue which is positive and that the eigenvalues of G are ordered such that

$$|\sigma_n| \leq |\sigma_{n-1}| \leq \cdots \leq |\sigma_3| \leq |\sigma_2| < \sigma_1 .$$

We let $\underline{x}_i$ be the eigenvector associated with the eigenvalue σ_i, i.e., $G\underline{x}_i = \sigma_i \underline{x}_i$.

Unless the contrary is explicitly stated, we also assume that σ_2 is real and positive and that $\sigma_2 > |\sigma_i|$ for $i \geq 3$.

In this chapter, we are concerned with the problem of solving the homogeneous equation

$$G\underline{x} = \sigma\underline{x} \tag{2.1}$$

for the dominant eigenvalue σ_1 and its corresponding eigenvector $\underline{x}_1$.

2. The Power Method

One may iteratively solve the eigenvalue problem (2.1) using the well-known power method. Given the real initial vector $\underline{x}(0)$ and eigenvalue $\sigma(0)$, the power method generates successive estimates for the eigenvector $\underline{x}_1$ and eigenvalue σ_1 by the process

$$\begin{cases} \underline{v}(k) = \dfrac{G}{\sigma(k-1)}\,\underline{x}(k-1) \\[2ex] \sigma(k) = \sigma(k-1)\,\dfrac{[\underline{v}(k),\underline{v}(k)]}{[\underline{v}(k),\underline{x}(k-1)]} \\[2ex] \underline{x}(k) = \underline{v}(k) \ , \end{cases} \tag{2.2}$$

where $[\underline{r},\underline{s}]$ denotes the scalar product of the vector $\underline{r}$ with the vector $\underline{s}$, i.e., $[\underline{r},\underline{s}] = \underline{r}^*\underline{s}$ and $\underline{r}^*$ is the complex conjugate transpose of the vector $\underline{r}$. The integer k in (2.2) is the iteration index number.

There are many ways by which the eigenvalue may be estimated in the power method. The eigenvalue estimate $\sigma(k)$ in (2.2) is obtained by the so-called modified Rayleigh quotient [Bilodeau and Hageman (1957)]. Other techniques which may be used to estimate the eigenvalue are the Rayleigh quotient, the component sum, and the single component techniques. Unless the matrix G is symmetric[1], it usually makes very little difference which technique is used to estimate the eigenvalue.

For the power method of iteration, the eigenvector is more crucial and more evasive than the eigenvalue. Intuitively, this may be seen by considering the eigenvalue problem (2.1). Given the eigenvalue σ_1, it is still a difficult task to determine $\underline{x}_1$; whereas, given the eigenvector $\underline{x}_1$, it is easy to calculate σ_1. Thus, for the most part, we shall concentrate on the convergence of the eigenvector.

For any iterative process, the answers to several questions must be considered. For example,

1. Does the iterative process converge?

2. If the process is convergent, how fast or at what rate does it converge?

3. What practical criterion may be used to terminate the iterative process?

In this chapter, we shall be concerned with answers to the first two questions. The third question will be discussed in a later chapter.

---

[1] If the matrix G is symmetric, the Rayleigh quotient and the modified Rayleigh quotient have certain advantages over the other techniques.

Since the matrix G has a dominant eigenvalue, it follows [Faddeev and Faddeeva (1963)] that the power iterative method (2.2) is convergent, i.e.,

$$\lim_{k\to\infty} \sigma(k) = \sigma_1 \quad \text{and} \quad \lim_{k\to\infty} \underline{x}(k) = \underline{x}_1 .$$

The rate of convergence of the power method depends primarily on how well separated the dominant eigenvalue σ_1 is from the other eigenvalues of G. To see why this is true, let us first assume[1] that the eigenvectors of G span $V_n(C)$. Thus, the eigenvector estimate $\underline{x}(k_1)$ after k_1 iterations may be written as

$$\underline{x}(k_1) = \underline{x}_1 + \sum_{i=2}^{n} c_i \underline{x}_i , \tag{2.3}$$

where the c_i are scalars. The corresponding error vector $\underline{E}(k_1)$ for iteration k_1 can be expressed by

$$\underline{E}(k_1) \equiv \underline{x}(k_1) - \underline{x}_1 = \sum_{i=2}^{n} c_i \underline{x}_i .$$

For iteration $(k_1 + 1)$, we have

$$\underline{x}(k_1 + 1) = \frac{G}{\sigma(k_1)} \underline{x}(k_1) = \frac{\sigma_1}{\sigma(k_1)} \underline{x}_1 + \sum_{i=2}^{n} \left(\frac{\sigma_i}{\sigma(k_1)}\right) c_i \underline{x}_i . \tag{2.4}$$

If we now assume that k_1 is large enough so that the eigenvalue estimates $\sigma(k_1 + r)$, $r \geq 0$, are sufficiently close to σ_1, then for iteration $(k_1 + r)$, we have

---

[1] This assumption will be abandoned only in the last section of this chapter.

(2.5) $$\underline{x}(k_1 + r) \approx \left(\frac{G}{\sigma_1}\right)^r \underline{x}(k_1) = \underline{x}_1 + \sum_{i=2}^{n} \left(\frac{\sigma_i}{\sigma_1}\right)^r c_i \underline{x}_i$$

and

(2.6) $$\underline{E}(k_1 + r) \approx \sum_{i=2}^{n} \left(\frac{\sigma_i}{\sigma_1}\right)^r c_i \underline{x}_i \ .$$

Thus, the rate at which the error vector $\underline{E}(k)$ approaches the null vector or equivalently the rate at which $\underline{x}(k)$ approaches $\underline{x}_1$ depends on how well separated the dominant eigenvalue σ_1 is from the other eigenvalues of G.

If the dominance ratio d of the matrix G is defined by

(2.7) $$d \equiv \max_{i \neq 1} \frac{|\sigma_i|}{|\sigma_1|} = \frac{|\sigma_2|}{|\sigma_1|} \ ,$$

then the most slowly decaying contribution to the error vector is multiplied by a scalar of modulus d each iteration. Thus, d may be taken as the average reduction factor per iteration for successive error vectors. We define the average rate of convergence R for the power method as

(2.8) $$R \equiv -\ln d \ .$$

Roughly speaking, the reciprocal of R is a measure of the number of iterations required to reduce the initial error vector by a factor e, where e is the base of the natural logarithms. Thus, a natural criterion for the comparison of different iterative methods is the size of their respective rates of convergence. For a more detailed discussion on convergence rates, see Varga (1962), page 62.

In the next section we shall describe the Chebyshev polynomial extrapolation method which often may be used to accelerate the convergence rate of the basic power method.

3. The Chebyshev Polynomial Extrapolation Method

From Eq. (2.5) we see that the performance of r power iterations results in the most slowly decaying contribution to the error vector being multiplied by a factor of d^r. We note that these r power iterations correspond to applying the matrix operator $\left(\frac{G}{\sigma_1}\right)^r$ to the vector $\underline{x}(k_1)$. Now if a r-th degree matrix polynomial[1] $Q_r\left(\frac{G}{\sigma_1}\right)$ were used to operate on $\underline{x}(k_1)$, we could express[2] $\underline{x}(k_1 + r)$ as

$$(2.9) \qquad \underline{x}(k_1 + r) = Q_r\left(\frac{G}{\sigma_1}\right)\underline{x}(k_1) = Q_r(1)\underline{x}_1 + \sum_{i=2} Q_r\left(\frac{\sigma_i}{\sigma_1}\right)c_i\underline{x}_i \ .$$

Hence, if we could choose the polynomial $Q_r(y)$ such that $Q_r(1) = 1$ and $\sum_{i=2} Q_r\left(\frac{\sigma_i}{\sigma_1}\right)c_i\underline{x}_i = \underline{0}$, then we would have $\underline{x}(k_1 + r) = \underline{x}_1$. Even if such a polynomial existed, it would be a function of the c_i, $\underline{x}_i$, and σ_i, which generally are not known for all i. Therefore, such a special polynomial is usually out of the question.

Suppose, however, that the eigenvalues σ_i of G are real and satisfy $b \leq \sigma_i/\sigma_1 \leq d$ for $i \geq 2$.[3] Then we can try to choose for $Q_r(y)$ that polynomial $P_r(y)$ having the least maximum modulus over the range $b \leq y \leq d$ and such that

---

[1] If $Q_r(y) = \sum_{k=0}^{r} b_k y^k$ is of polynomial of degree r in y, then the matrix polynomial $Q_r(B)$ in the matrix B is defined as $Q_r(B) = \sum_{k=0}^{r} b_k B^k$.

[2] We are still assuming that the eigenvectors $\underline{x}_i$ of G span $V_n(C)$.

[3] In keeping with the assumption made in Section 1 of this chapter, we assume that $d > |b|$.

$P_r(1) = 1$. Such a polynomial exists [Flanders and Shortley (1950)] and can be given explicitly in terms of Chebyshev polynomials by

$$P_r(y) \equiv \frac{T_r\left(\frac{2y - d - b}{d - b}\right)}{T_r\left(\frac{2 - d - b}{d - b}\right)} , \tag{2.10}$$

where $T_r(w)$ is the Chebyshev polynomial of degree r. For $r > 0$, $T_r(w) \equiv \cos[r\cos^{-1}w]$ if $|w| \leq 1$ and $T_r(w) \equiv \cosh[r\cosh^{-1}w]$ if $|w| \geq 1$.[1]

With $Q_r(y) = P_r(y)$, the polynomial method of (2.9) is called the Chebyshev polynomial method and the matrix $\left(\frac{G}{\sigma_1}\right)$ is called the <u>argument matrix</u>.

The well known recurrence relation for Chebyshev polynomials

$$T_r(w) = 2wT_{r-1}(w) - T_{r-2}(w) , \quad r \geq 2 , \tag{2.11}$$

where $T_0(w) = 1$ and $T_1(w) = w$, enables one to successively generate the polynomials $P_r(y)$ in a straightforward way [see, for example, Hageman (1963) p. 27]. Starting with $\underline{x}(k_1)$, one may generate successively

$$\begin{aligned} \underline{x}(k_1 + 1) &= P_1\left(\frac{G}{\sigma_1}\right)\underline{x}(k_1) \\ \underline{x}(k_1 + 2) &= P_2\left(\frac{G}{\sigma_1}\right)\underline{x}(k_1) \\ &\vdots \\ \underline{x}(k_1 + r) &= P_r\left(\frac{G}{\sigma_1}\right)\underline{x}(k_1) \end{aligned}$$

[1] If w is any complex number, then $T_r(w)$ may be expressed [Forsythe and Wasow (1960)] as

$$T_r(w) \equiv 1/2\left[(w + \sqrt{w^2 - 1})^r + (w - \sqrt{w^2 - 1})^r\right] .$$

using the procedure

$$
(2.12)\quad \begin{cases}
\underline{v}(k_1 + t) = \dfrac{G}{\sigma(k_1 + t - 1)}\, \underline{x}(k_1 + t - 1) \\[2ex]
\underline{x}(k_1 + t) = \underline{x}(k_1 + t - 1) + \alpha_{k_1+t}[\underline{v}(k_1 + t) - \underline{x}(k_1 + t - 1)] \\[1ex]
\qquad\qquad\qquad + \beta_{k_1+t}[\underline{x}(k_1 + t - 1) - \underline{x}(k_1 + t - 2)] \\[2ex]
\sigma(k_1 + t) = \sigma(k_1 + t - 1)\, \dfrac{[\underline{v}(k_1 + t)\, ,\, \underline{v}(k_1 + t)]}{[\underline{v}(k_1 + t), \underline{x}(k_1 + t - 1)]}\, ,
\end{cases}
$$

for $t=1,2,3,\ldots$ (The σ calculation is included in (2.12) to take into account the fact that $\sigma(k_1)$ is not exactly equal to σ_1.) α_{k_1+t} and β_{k_1+t} are functions of d, b, and t and are given by

$$
(2.13)\quad \begin{cases}
\alpha_{k_1+1} = \dfrac{2}{2 - d - b}\ ;\ \beta_{k_1+1} = 0 \text{ and for } t \geq 2 \\[2ex]
\alpha_{k_1+t} = \dfrac{4}{d - b}\, \dfrac{T_{t-1}\left(\dfrac{2 - d - b}{d - b}\right)}{T_t\left(\dfrac{2 - d - b}{d - b}\right)}\ ;\ \beta_{k_1+t} = \dfrac{T_{t-2}\left(\dfrac{2 - d - b}{d - b}\right)}{T_t\left(\dfrac{2 - d - b}{d - b}\right)}\ .
\end{cases}
$$

Since

$$
(2.14)\qquad \max_{b \leq y \leq d} |P_r(y)| = P_r(d) = \frac{1}{T_r\left(\dfrac{2 - d - b}{d - b}\right)}\ ,
$$

we see that the most slowly decaying contributions to the error vector are multiplied by a factor of modulus $P_r(d)$ in r iterations. For d close to unity, the Chebyshev polynomial method of iteration is an order of magnitude faster than the power method. Table 2.1 shows as a function of d the gain in speed of convergence one may obtain using the Chebyshev polynomial method of iteration compared to the power method. For Table 2.1, b is assumed to be zero.

d	1 ITERATION		3 ITERATIONS		5 ITERATIONS		10 ITERATIONS	
	$(d)^1$	$P_1(d)$	$(d)^3$	$P_3(d)$	$(d)^5$	$P_5(d)$	$(d)^{10}$	$P_{10}(d)$
.6	.6	.429	.216	.023	.078	.001	.006	----
.7	.7	.538	.343	.049	.168	.004	.028	----
.8	.8	.667	.512	.111	.328	.016	.108	----
.9	.9	.818	.729	.276	.590	.076	.349	.003
.95	.95	.905	.857	.481	.774	.204	.599	.021
.97	.97	.942	.913	.624	.859	.337	.737	.060
.985	.985	.970	.956	.778	.927	.539	.860	.170
.99	.99	.980	.970	.843	.951	.647	.904	.266
.992	.992	.984	.976	.871	.961	.700	.923	.325
.995	.995	.990	.985	.917	.975	.792	.951	.457
.998	.998	.996	.994	.965	.990	.908	.980	.700

TABLE 2.1

From (2.14) we see that $[P_r(d)]^{1/r}$ is the average reduction factor per iteration for successive error vectors of the Chebyshev polynomial method. Thus, the average rate of convergence for r iterations of the Chebyshev polynomials method is defined [Varga (1962), page 134] as

$$R_r[P_r] \equiv -\ln[P_r(d)]^{1/r} \, . \tag{2.15}$$

From Eq. (2.14), it follows that $R_r[P_r]$ increases monotonically with r and that [Varga (1962), page 139]

$$(2.16) \qquad R_{\infty}[P_r] \equiv \lim_{r \to \infty} R_r[P_r] = \cosh^{-1}\left(\frac{2 - d - b}{d - b}\right) .$$

In a later section, we shall discuss convergence rates in more detail.

Thus far, we have assumed that the eigenvalue bounds d and b are known. This, of course, is not a realistic assumption. If the estimates for d and b are denoted by d_0 and b_0, then we shall take $P_{r,\sigma_0}(y)$ to be the Chebyshev polynomial of degree r in which d_0 and b_0 are used as estimates for d and b. From Eq. (2.10), we see that $P_{r,d_0}(y)$ may be written as

$$(2.17) \qquad P_{r,d_0}(y) = \frac{T_r\left(\dfrac{2y - d_0 - b_0}{d_0 - b_0}\right)}{T_r\left(\dfrac{2y - d_0 - b_0}{d_0 - b_0}\right)}$$

and hence

$$(2.18) \qquad \max_{b \le y \le d} |P_{r,d_0}(y)| = \frac{\max\limits_{b \le y \le d} \left|T_r\left(\dfrac{2y - d_0 - b_0}{d_0 - b_0}\right)\right|}{T_r\left(\dfrac{2 - d_0 - b_0}{d_0 - b_0}\right)} .$$

From the min-max property of Chebyshev polynomials or by directly comparing (2.14), and (2.18), we have

$$\max_{b \le y \le d} |P_r(y)| \le \max_{b \le y \le d} |P_{r,d_0}(y)|$$

with equality only if $d_0 = d$ and $b_0 = b$.

To illustrate how the effectiveness of the Chebyshev polynomial method depends on the estimate of d, let us consider a matrix G for which

$d = .9$ and $b = 0$. If the estimates for d and b were correct, i.e., $d_0 = d = .9$ and $b_0 = b = 0$, then from Table 2.1 we have $P_5(d) = .076$; whereas, with $d_0 = .8$ and $b_0 = 0$, we have from (2.18) that $\max_{b \le y \le d} |P_{5,d_0}(y)| = P_{5,d_0}(d) \approx .25$. Thus, in 5 iterations the most slowly converging contributions to the error vector are multiplied by a factor of modulus .076 in the optimum parameter case as compared to .25 in the non-optimum parameter case. Hence, the use of non-optimum values for d and b can result in a sizable reduction in the convergence rate of the Chebyshev method of iteration. Fortunately, as we shall see in the next chapter, practical numerical means exist for estimating these unknown constants.

In addition to the basic assumption that G has a dominant eigenvalue σ_1, we have assumed, thus far, that the eigenvalues of G are real and that the eigenvectors of G span the associated vector space of G.[1] In the next section we shall relax the assumption that the eigenvalues of G be real.

4. Complex Eigenvalues

In this section we will again assume that the eigenvectors $\{\underline{x}_i\}_{i=1}^{i=n}$ of G span the vector space $V_n(C)$ but the assumption that all the eigenvalues of G are real will be relaxed. We shall assume only that the dominant eigenvalue σ_1 is real and positive.

We will present two approaches which, hopefully, will illustrate the effect of complex eigenvalues on the Chebyshev polynomial method of iteration. The first approach will be to show the effect of complex eigenvalues on the convergence rate when the Chebyshev polynomial of Eq. (2.10) is applied. We

[1]We have also assumed that σ_1 and σ_2 are positive and that $\sigma_2 > |\sigma_i|$ for $i \geq 3$. These assumptions, however, were made merely for reasons of simplicity and are not restricting.

note that the Chebyshev polynomial given in Eq. (2.10) is based on the assumption that the eigenvalues of G are real. The second approach will be to change the argument of the Chebyshev polynomial so that the min-max property of these polynomials will be valid over part of the complex plane.[1]

A. Complex Eigenvalues and the Real Domain Chebyshev Polynomial

Let the eigenvalues of G be denoted by $\left\{\sigma_i\right\}_{i=1}^{i=n}$. As before, we take the dominant eigenvalue σ_1 to be real and positive but now we assume only that the quantities $\left\{\frac{\sigma_i}{\sigma_1}\right\}_{i=2}^{i=n}$ are contained in a connected region D in the complex plane. The dominance ratio is again given by $d = |\sigma_2|/\sigma_1$.

Now suppose that the Chebyshev polynomial defined by Eq. (2.10) is used in the polynomial method of (2.9). If d_0 and b_0 are used as estimates for d and b in (2.10) then, as in Eq. (2.17), we let

$$P_{r,d_0}(z) = \frac{T_r\left(\frac{2z - d_0 - b_0}{d_0 - b_0}\right)}{T_r\left(\frac{2 - d_0 - b_0}{d_0 - b_0}\right)} . \tag{2.19}$$

Thus, from Eq. (2.9) we see that in r iterations the most slowly decaying contributions to the error vector are multiplied by at most a factor of modulus $f_r(D)$, where

$$f_r(D) \equiv \max_{z \in D} |P_{r,d_0}(z)| = \frac{\max_{z\in D}\left|T_r\left(\frac{2z - d_0 - b_0}{d_0 - b_0}\right)\right|}{T_r\left(\frac{2 - d_0 - b_0}{d_0 - b_0}\right)} . \tag{2.20}$$

[1]The effect of complex eigenvalues on Chebyshev extrapolation is also discussed by Wachspress (1966) and Wrigley (1963).

The average reduction factor per iteration is then

$$F_r(D) \equiv [f_r(D)]^{1/r} . \tag{2.21}$$

We now wish to determine how $F_r(D)$ is affected by the region D.

If the eigenvalues of G are real and if d_0 and b_0 satisfy $d_0 \geq d$ and $b_0 \leq b$, then the region D may be chosen to be the closed interval $[b_0,d_0]$ and for this case we have

$$\left\{F_r([b_0,d_0])\right\}^r = \max_{z\in[b_0,d_0]} |P_{r,d_0}(z)| = \frac{1}{T_r\left(\frac{2 - d_0 - b_0}{d_0 - b_0}\right)} . \tag{2.22}$$

Without more knowledge concerning the eigenvalues of G, $F_r([b_0,d_0])$ is the smallest average reduction factor which can be achieved by the Chebyshev polynomial method of iteration. Thus, in seeing how $F_r(D)$ is affected by the region D, we shall use $F_r([b_0,d_0])$ as the norm. In what follows, we shall denote $F_r([b_0,d_0])$ simply by F_r.

Let $D_r(c)$ be the set of points in the complex z plane such that the inequality $|P_{r,d_0}(z)| \leq \left\{cF_r\right\}^r$ is satisfied, i.e.,

$$D_r(c) \equiv \left\{z: |P_{r,d_0}(z)| \leq \left\{cF_r\right\}^r\right\}. \tag{2.23}$$

Thus, if all eigenvalues of $\left(\frac{G}{\sigma_1}\right)$ except unity, are contained in $D_r(c)$, then the average reduction factor per iteration achieved by the Chebyshev polynomial method will be less than or equal to cF_r. We are only interested in c over the range $1 \leq c \leq 1/F_r$. For if $c \geq 1/F_r$, the Chebyshev method of iteration is divergent. If $c < 1$, the set $D_r(c)$ consists of r separated regions which are centered about the r real zeros of the polynomial $T_r\left(\frac{2z - d_0 - b_0}{d_0 - b_0}\right)$ and thus not generally of practical interest.

Obviously, if c is fixed, then $D_r(c)$ is a function of r. If r = 1, the region $D_1(c)$ consists of all points on or interior to a circle with center at $\left(\frac{d_0 + b_0}{2}, 0\right)$ and radius $c\left(\frac{d_0 - b_0}{2}\right)$. See Figure 2.1.

The region $D_2(c)$ consists of all points on or interior to the ovals of Cassini. A proof of this together with the region $D_2(c)$ for an arbitrary c is given in Appendix A. The regions $D_2(c)$ for c = 1 and $c = 1/F_2$ are given in Figure 2.2. Thus, if we choose to cyclically apply[1] the polynomial $P_{2,d_0}(z)$ and if all the eigenvalues of $\left(\frac{G}{\sigma_1}\right)$ except unity, were contained in $D_2(c)$, then the average reduction factor per iteration would be no greater than cF_2.

In the limit as r approaches infinity, the region $D_\infty(c)$ consists of all points on or interior to the ellipse

$$(2.24) \qquad \frac{\left[x - \left(\frac{d_0 + b_0}{2}\right)\right]^2}{\left[\left(\frac{d_0 - b_0}{4}\right)\left(c + \frac{1}{c}\right)\right]^2} + \frac{y^2}{\left[\left(\frac{d_0 - b_0}{4}\right)\left(c - \frac{1}{c}\right)\right]^2} = 1 .$$

A proof of this is given in Appendix A. The regions $D_\infty(c)$ for c = 1 and[2] $c = 1/F_\infty$ are given in Figure 2.3. Note that $D_\infty(1)$ is simply the line segment $b_0 < z < d_0$.

We now shall give a more basic approach to the complex eigenvalue problem.

---

[1]Instead of letting r tend to infinity.

[2]The limit of F_r as $r \to \infty$ can be expressed as

$$F_\infty = \frac{d_0 - b_0}{2 - (d_0 + b_0) + 2\sqrt{(1 - d_0)(1 - b_0)}} .$$

This will be discussed in more detail in Section 5 of this chapter.

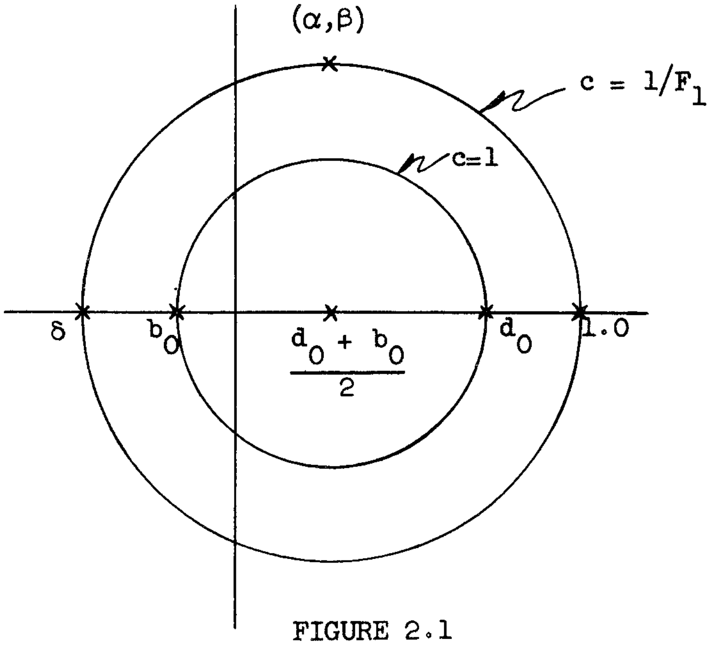


FIGURE 2.1
D_1 Region

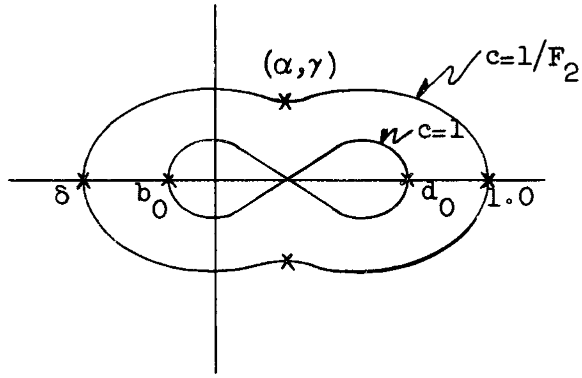


FIGURE 2.2
D_2 Region

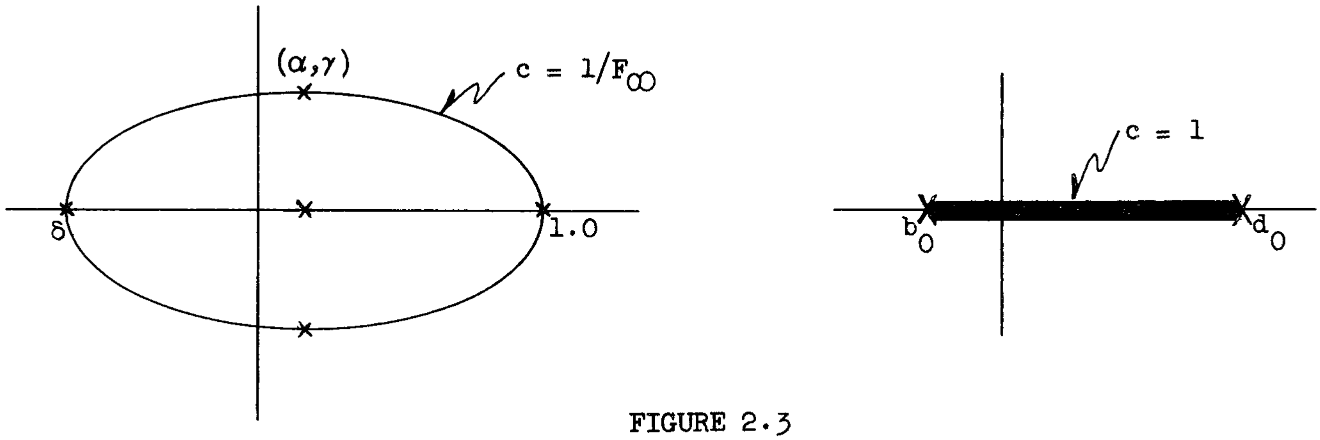


FIGURE 2.3
D_∞ Region

$$\alpha = \frac{d_0 + b_0}{2} \qquad \gamma = \sqrt{(1 - d_0)(1 - b_0)}$$

$$\beta = 1 - \left(\frac{d_0 + b_0}{2}\right) \qquad \delta = -1 + (d_0 + b_0)$$

B. The Complex Domain Chebyshev Polynomial

Suppose now that the quantities $\left\{\sigma_i/\sigma_1\right\}_{i=2}^{i=n}$ are contained in the ellipse (see Figure 2.4)

$$\frac{\left[x - \left(\frac{d+b}{2}\right)\right]^2}{\left(\frac{d-b}{2}\right)^2} + \frac{y^2}{\epsilon^2} = 1 , \tag{2.25}$$

where $0 < \epsilon < \frac{d-b}{2}$. Thus, the real eigenvalue premise used in obtaining

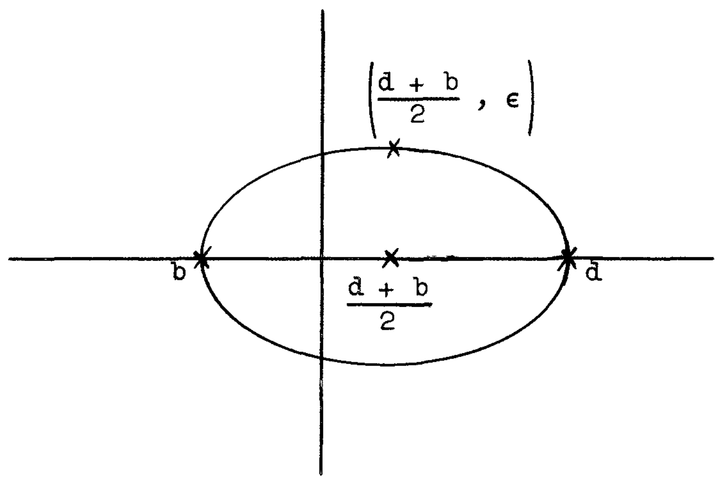


Figure 2.4

the polynomial (2.10) is not valid. For the complex case one would like to choose for $Q_r(y)$ in Eq. (2.9) that polynomial $\tilde{P}_r(y)$ having the least maximum modulus over the ellipse (2.25) and its interior. Clayton (1963) has shown that such a polynomial exists and that it is unique, real, and can be expressed in terms of Chebyshev polynomials by

$$(2.26)\qquad \tilde{P}_r(y) \equiv \frac{T_r\left(\frac{2y - d - b}{[(d - b)^2 - 4\epsilon^2]^{1/2}}\right)}{T_r\left(\frac{2 - d - b}{[(d - b)^2 - 4\epsilon^2]^{1/2}}\right)}$$

where the $T_r(w)$ are again Chebyshev polynomials of degree r. When $\epsilon = 0$, $\tilde{P}_r(y)$ reduces to $P_r(y)$. We shall refer to the polynomial method of (2.9) as the complex Chebyshev polynomial method when $Q_r(y) = \tilde{P}_r(y)$.

The vectors $\underline{x}(k_1 + r) = \tilde{P}_r\left(\frac{G}{\sigma_1}\right)\underline{x}(k_1)$ for the complex Chebyshev polynomial method may be generated successively using the same procedure (2.12) as described for the real Chebyshev case. For the complex case, however, the parameters α and β are given by

$$\alpha_{k_1+1} = \frac{2}{2 - d - b} \ ; \ \beta_{k_1+1} = 0 \text{ and for } t \geq 2$$

$$\alpha_{k_1+t} = \frac{4}{[(d-b)^2 - 4\epsilon^2]^{1/2}} \frac{T_{t-1}\left(\frac{2 - d - b}{[(d-b)^2 - 4\epsilon^2]^{1/2}}\right)}{T_t\left(\frac{2 - d - b}{[(d-b)^2 - 4\epsilon^2]^{1/2}}\right)} \ ; \ \beta_{k_1+t} = \frac{T_{t-2}\left(\frac{2 - d - b}{[(d-b)^2 - 4\epsilon^2]^{1/2}}\right)}{T_r\left(\frac{2 - d - b}{[(d-b)^2 - 4\epsilon^2]^{1/2}}\right)} .$$

If we let D denote the ellipse (2.25) and its interior, then one may easily show that

$$(2.27)\qquad \max_{z\epsilon D}|\tilde{P}_r(z)| = \tilde{P}_r(d) = \frac{T_r\left(\frac{d - b}{[(d - b)^2 - 4\epsilon^2]^{1/2}}\right)}{T_r\left(\frac{2 - d - b}{[(d - b)^2 - 4\epsilon^2]^{1/2}}\right)}$$

so that $[\tilde{P}_r(d)]^{1/r}$ is the average reduction factor per iteration for successive error vectors of the complex Chebyshev polynomial method.

As before, we define the average rate of convergence for r iterations of the complex Chebyshev method to be

$$(2.28) \qquad R_r[\tilde{P}_r] \equiv -\ln[\tilde{P}_r(d)]^{1/r} .$$

From Eq. (2.27) one may show that $R_r[\tilde{P}_r]$ increases monotonically with r and that

$$(2.29) \quad R_\infty[\tilde{P}_r) \equiv \lim_{r\to\infty} R_r[\tilde{P}_r] = \cosh^{-1}\left(\frac{2-d-b}{[(d-b)^2-4\epsilon^2]^{1/2}}\right) - \cosh^{-1}\left(\frac{d-b}{[(d-b)^2-4\epsilon^2]^{1/2}}\right).$$

Since $\tilde{P}_r(y)$ belongs to the set of polynomials from which $P_r(y)$ was chosen, we must have that $P_r(d) \le \tilde{P}_r(d)$ and hence $R_r[\tilde{P}_r] \le R_r[P_r]$ with strict inequality for all $r \ge 2$. Thus, the complex Chebyshev polynomial method does not achieve as great an improvement over the straight power method as does the real Chebyshev polynomial method. In the next section we will compare the quantities R, $R_\infty[P_r]$, and $R_\infty[\tilde{P}_r]$ when d is close to unity.

5. Rates of Convergence

As given previously, the average rate of convergence for the complex Chebyshev polynomial method increases monotonically with r to the limit $R_\infty[\tilde{P}_r]$, where

$$R_\infty[\tilde{P}_r] = \cosh^{-1}\left(\frac{2-d-b}{[(d-b)^2-4\epsilon^2]^{1/2}}\right) - \cosh^{-1}\left(\frac{d-b}{[(d-b)^2-4\epsilon^2]^{1/2}}\right) .$$

Since $\cosh^{-1}(y) = \ln[y + \sqrt{y^2-1}]$, we have

$$R_\infty[\tilde{P}_r] = \ln\left\{\frac{(2-d-b) + [(2-d-b)^2 - (d-b)^2 + 4\epsilon^2]^{1/2}}{d-b+2\epsilon}\right\}$$

or equivalently

$$(2.30) \qquad R_\infty[\tilde{P}_r] = -\ln\left\{\frac{d - b + 2\epsilon}{2 - d - b + 2[(1 - d)(1 - b) + \epsilon^2]^{1/2}}\right\} .$$

Similarly, for the real Chebyshev polynomial method we have

$$(2.31) \qquad R_\infty[P_r] = -\ln\left\{\frac{d - b}{2 - d - b + 2\,[(1 - d)(1 - b)]^{1/2}}\right\} .$$

We recall that the quantities d, b, and ϵ are assumed to satisfy

$$(2.32) \qquad d < 1 \ , \ |b| < d \ , \text{ and } \epsilon < \frac{d - b}{2} .$$

The average rate of convergence for the power method does not depend on r and from Eq. (2.8) is given by

$$(2.33) \qquad R = -\ln d .$$

We will now compare these convergence rates when d is near unity or equivalently when δ, where

$$(2.34) \qquad \delta = 1 - d ,$$

is near zero. Since $-\ln y = (1 - y) + \frac{(1 - y)^2}{2} + \frac{(1 - y)^3}{3} + \ldots$ for $0 < y \leq 1$, we may write Eq. (2.33) as

$$(2.35) \qquad R = \delta + \frac{\delta^2}{2} + \frac{\delta^3}{3} + \ldots$$

Thus, for small δ, a good approximation for the rate of convergence of the power method is

$$R \approx \delta \quad (2.36)$$

Similarly, $R_\infty[P_r]$ may be expressed as

$$R_\infty[P_r] = \delta[P_r] + \frac{\{\delta[P_r]\}^2}{3} + \frac{\{\delta[P_r]\}^3}{3} + \cdots \quad , \quad (2.37)$$

where

$$\delta[P_r] = 1 - \frac{d - b}{2 - d - b + 2[(1 - d)(1 - b)]^{1/2}} = 2\left[\frac{\delta + [\delta(1 - b)]^{1/2}}{\delta + (1 - b) + 2[\delta(1 - b)]^{1/2}}\right].$$

Since $1 - b > 0$, we may write $\delta[P_r]$ as

$$\delta[P_r] = \frac{2\sqrt{\frac{\delta}{1 - b}} + 2\left(\frac{\delta}{1 - b}\right)}{1 + 2\sqrt{\frac{\delta}{1 - b}} + \frac{\delta}{1 - b}} \quad . \quad (2.38)$$

In most practical applications b will range from 0 to -d so that $1 - b$ usually varies from 1 to 2. Thus, for small δ it is reasonable to assume that $\delta/(1 - b)$ is also small. Hence, for small δ, a good approximation for the rate of convergence of the real Chebyshev polynomial method is

$$R_\infty[P_r] \approx 2\sqrt{\frac{\delta}{1 - b}} \quad . \quad (2.39)$$

Note that as b varies from 0 to -d, $R_\infty[P_r]$ varies only from $2\sqrt{\delta}$ to $\sqrt{2} \cdot \sqrt{\delta}$. Thus $R_\infty[P_r]$ is not greatly affected by the value of b.

For the complex Chebyshev polynomial, $R_\infty[\tilde{P}_r]$ may be expressed as

$$(2.40)\qquad R_\infty[\tilde{P}_r] = \delta[\tilde{P}_r] + \frac{\{\delta[\tilde{P}_r]\}^2}{2} + \frac{\{\delta[\tilde{P}_r]\}^3}{3} + \cdots ,$$

where

$$\delta[\tilde{P}_r] = 1 - \frac{d - b + 2\epsilon}{2 - d - b + 2[(1-d)(1-b)+ \epsilon^2]^{1/2}} = 2\left[\frac{\delta - \epsilon + [\delta(1-b) + \epsilon^2]^{1/2}}{\delta + (1-b) + 2[\delta(1-b) + \epsilon^2]^{1/2}}\right].$$

If we let $K^2 \equiv \epsilon^2/[(1 - b)\delta]$, then $\delta[\tilde{P}_r]$ may be written as

$$(2.41)\qquad \delta[\tilde{P}_r] = \frac{2\sqrt{\frac{\delta}{1 - b}}\sqrt{1 + K^2} - K + 2\frac{\delta}{1 - b}}{1 + 2\sqrt{\frac{\delta}{1 - b}}\sqrt{1 + K^2} + \frac{\delta}{1 - b}} .$$

Now ϵ must satisfy $0 \le \epsilon \le \frac{d - b}{2}$ so that K must satisfy $0 \le K < \frac{1}{\sqrt{\delta}}\left[\frac{d - b}{2\sqrt{1 - b}}\right]$. Thus, for small δ, K may take on large values. For $\epsilon = 0$ $(K = 0)$, $\delta[\tilde{P}_r]$ is the same as $\delta[P_r]$ and thus for small δ

$$(2.42)\qquad R_\infty[\tilde{P}_r]_{\epsilon=0} \approx 2\sqrt{\frac{\delta}{1 - b}} .$$

For $\epsilon = \sqrt{\delta(1 - b)}$ $(K = 1)$, we have for small δ

$$(2.43)\qquad R_\infty[\tilde{P}_r]_\epsilon = \sqrt{\delta(1 - b)} \approx .828\sqrt{\frac{\delta}{1 - b}} .$$

Thus, as ϵ varies from 0 to only $\sqrt{\delta(1 - b)}$, the rate of convergence of the complex Chebyshev method varies by more than a factor of 2. We now shall see what happens to $R_\infty[\tilde{P}_r]$ as ϵ approaches $\frac{d - b}{2}$ or equivalently as K approaches

$$\frac{1}{\sqrt{\delta}}\,\frac{d - b}{2\sqrt{1 - b}} .$$

With $\epsilon > 0$ and using the fact that $\epsilon = K\sqrt{\delta(1 - b)}$, Eq. (2.41) may be written as

$$\text{(2.44)} \qquad \delta[\tilde{P}_r] = \frac{\delta}{\epsilon}\left\{\frac{2K[\sqrt{1 + K^2} - K] + 2K\sqrt{\frac{\delta}{1 - b}}}{1 + 2\sqrt{\frac{\delta}{1 - b}}\sqrt{1 + K^2} + \frac{\delta}{1 - b}}\right\} \equiv \frac{\delta}{\epsilon}\,\{f(K)\} .$$

For $K > 0$, $f(K)$ is an increasing function[1] of K and $f\left(\frac{1}{\sqrt{\delta}}\left[\frac{d - b}{2\sqrt{1 - b}}\right]\right) = \frac{d - b}{1 - b + \delta}$. Thus, $f(K) < 1$ for all $0 < K < \frac{1}{\sqrt{\delta}}\left[\frac{d - b}{2\sqrt{1 - b}}\right]$. Hence, for $\epsilon > 0$ and small δ we have

$$\text{(2.45)} \qquad R_\infty[\tilde{P}_r] < \frac{\delta}{\epsilon} .$$

As ϵ approaches $\frac{d - b}{2}$, we have from (2.44) and the above that

$$\text{(2.46)} \qquad R_\infty[\tilde{P}_r]_{\epsilon \to \frac{d-b}{2}} \approx \frac{\delta}{\epsilon}\,\frac{d - b}{1 - b + \delta} \approx \frac{2\delta}{1 - b + \delta} .$$

If $b = -d$, note that $R_\infty[\tilde{P}_r]$ approaches the convergence rate of the power method as ϵ approaches d. This agrees with the well-known result [Varga (1957)] that as the ellipse containing the normalized eigenvalues $\left\{\sigma_i/\sigma_1\right\}_{i=2}^{n}$ tends to a circle, the min-max polynomial defined by (2.26) tends to $\left(y - (\frac{d + b}{2})\right)^r$ or just y^r when $b = -d$.

---

[1]This may be easily shown from the derivative $f'(K)$.

From expressions (2.36), (2.39), and (2.45) we have for $\delta = 1 - d$ close to zero

(2.47) $$R_\infty[P_r] \approx \frac{2}{\sqrt{\delta(1-b)}} R$$

and for $\epsilon > 0$

(2.48) $$R_\infty[\tilde{P}_r] \lesssim \frac{R}{\epsilon} .$$

Table 2.2 indicates how the different convergence rates vary as a function of d, b, and ϵ.

d	R	$R_\infty[P_r]$			$R_\infty[\tilde{P}_r]$	
		b=0	b=-.1	b=-.3	b=-.1 ϵ=.166	b=-.3 ϵ=.4
.8	.223	.963	.911	.829	.621	.386
.9	.104	.654	.622	.569	.376	.213
.95	.051	.455	.433	.397	.225	.114
.99	.010	.200	.191	.176	.055	.024
.995	.002	.146	.136	.125	.012	.005

TABLE 2.2

VARIATION IN CONVERGENCE RATES WITH d, b, AND ϵ

Thus, when the eigenvalues are real and d close to unity, the real Chebyshev polynomial method is an order of magnitude faster than the power method. For the complex eigenvalue case, the complex Chebyshev polynomial method is likely to achieve a much smaller, though still welcome, increase in convergence rate over the power method.

We remark that the comparisons given above are based on R_∞ and not R_r. As mentioned previously, R_r increases monotonically with r and is bounded by $R_1 \le R_r \le R_\infty$. For r = 1 and for d close to unity we have

$$(2.49)\qquad \begin{cases} R_1[P_1] \approx \dfrac{2}{1 - b + \delta} R \quad \text{and} \\ \\ R_1[\tilde{P}_1] \approx R_1[P_1] \quad . \end{cases}$$

In the next section we will discuss the case when the eigenvectors do not span the associated vector space of G.

6. An Incomplete Set of Eigenvectors

In the previous material we have assumed that the set of eigenvectors of G spans the associated vector space of G. In this section we abandon this requirement. Before proceeding, we first give a preliminary discussion on the concept of principal vectors.

A. Principal Vectors

As used in this report, a vector is simply an ordered collection of n complex numbers. The totality of all such vectors with n elements or components is called the n-dimensional vector space over the complex number field and is denoted by $V_n(C)$. Since the nxn matrix G with complex elements operating on a vector $\underline{x}$ in $V_n(C)$ merely transforms $\underline{x}$ into another vector $\underline{y}$ in $V_n(C)$, we say that $V_n(C)$ is the vector space associated with the matrix A.

The set of vectors $\left\{\underline{y}_i\right\}_{i=1}^{i=t}$ are said to span the vector space $V_n(C)$ if every vector in $V_n(C)$ can be written as a linear combination of $\underline{y}_i$. If the set $\left\{\underline{y}_i\right\}_{i=1}^{i=t}$ spans $V_n(C)$, then [Perlis (1952)] $t \geq n$ and the set $\left\{\underline{y}_i\right\}_{i=1}^{i=t}$ contains precisely n linearly independent vectors, i.e., any set of n + 1 vectors from $\left\{\underline{y}_i\right\}_{i=1}^{i=t}$ is dependent.

If the set of n vectors $\left\{\underline{y}_i\right\}_{i=1}^{i=n}$ is linearly independent, then this set spans $V_n(C)$ and is said to form a basis for $V_n(C)$. Thus, any set of n linearly independent vectors forms a basis for $V_n(C)$ and hence also spans $V_n(C)$. We now wish to define a basis for $V_n(C)$ in terms of the eigenvectors and principal vectors of the matrix G.

An nxn matrix G with complex elements has precisely n eigenvalues associated with it. These eigenvalues are defined to be the n roots of the characteristic equation[1]

(2.50) $$|G - zI| = z^n + h_{n-1}z^{n-1} + \dots + h_1 z + h_0 = 0 .$$

The roots of (2.50) are not necessarily distinct. The number of roots to (2.50) which have the same value is called the multiplicity of that root or eigenvalue. For what follows, let the eigenvalues of G be denoted by

(2.51) $$(\sigma_1)_{m_1}, (\sigma_2)_{m_2}, \dots, (\sigma_i)_{m_i}, \dots, (\sigma_t)_{m_t} ,$$

[1]By $|G - zI|$, we mean the determinant of the matrix $(G - zI)$.

where $\sigma_1, \sigma_2, \cdots, \sigma_t$ are all distinct and m_i is the multiplicity of the eigenvalue σ_i and where $\sum_{i=1}^{t} m_i = n$.

With each eigenvalue σ_i of G we may associate at least one nonzero vector $\underline{x}_i$ which satisfies the homogeneous equation

$$G\underline{x}_i = \sigma_i \underline{x}_i \tag{2.52}$$

or equivalently

$$(G - \sigma_i I)\underline{x}_i = \underline{0} \ . \tag{2.53}$$

The existence of at least one nonzero vector $\underline{x}_i$ is assured since $|G - \lambda I| = 0$. Any nonzero vector which satisfies (2.52) is called an eigenvector of G corresponding to the eigenvalue σ_i. From the set of eigenvectors for G, we would like to pick n linearly independent vectors to form a basis for $V_n(C)$. But, as we shall see, this is not always possible.

If the matrix G is normal[1], then it is known [Perlis (1952)] that it is possible to find a basis for $V_n(C)$ consisting of eigenvectors of G. (In fact, one may choose the vectors of this basis to be mutually orthogonal.) If the matrix G is not normal, it may not be possible[2] to find a basis for $V_n(C)$ from the set of eigenvectors of G. However, it is always possible to find a basis for $V_n(C)$ from the set of principal vectors of G.

---

[1]The matrix G is normal if $G^* G = GG^*$, where G^* is the conjugate transpose of G. Note that all Hermitian, skew-Hermitian, real symmetric, and real skew matrices are normal.

[2]One need only consider the matrix $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ to show this.

Any nonzero vector $\underline{y}_i$ which satisfies

$$(G - \sigma_i I)^p \underline{y}_i = \underline{0} \tag{2.54}$$

but for which

$$(G - \sigma_i I)^{p-1} \underline{y}_i \neq \underline{0} \tag{2.55}$$

is called [Householder (1953), page 32] a principal vector of grade p corresponding to the eigenvalue σ_i. Note that the set of eigenvectors is included in the set of principal vectors since eigenvectors are principal vectors of grade 1.

The following theorem is a restatement of results given in sections 57 and 58 of a book by Halmos (1957).

Theorem 2.1: For each of the distinct eigenvalues σ_i of (2.51), there exists positive integers $q, p_1, p_2, \ldots, p_q$ and nonzero vectors $\underline{y}_1, \underline{y}_2, \ldots, \underline{y}_q$ such that the m_i vectors

$$\left\{ \begin{array}{l} \underline{y}_1, (G - \sigma_i I)\underline{y}_1, \ldots, (G - \sigma_i I)^{p_1 - 1} \underline{y}_1 \\ \underline{y}_2, (G - \sigma_i I)\underline{y}_2, \ldots, (G - \sigma_i I)^{p_2 - 1} \underline{y}_2 \\ \vdots \qquad \vdots \qquad \vdots \\ \underline{y}_q, (G - \sigma_i I)\underline{y}_q, \ldots, (G - \sigma_i I)^{p_q - 1} \underline{y}_q \end{array} \right. \tag{2.56}$$

are linearly independent. Moreover,

$$(G - \sigma_i I)^{p_1} \underline{y}_1, = (G - \sigma_i I)^{p_2} \underline{y}_2 = \cdots = (G - \sigma_i I)^{p_q} \underline{y}_q = \underline{0}$$

and $p_1 + p_2 + \cdots + p_q = m_i$, where m_i is defined in (2.51).

It is easily seen that the m_i vectors given by (2.56) are principal vectors of G corresponding to σ_i and that the q vectors $(G - \sigma_i I)^{p_j - 1} \underline{y}_j$, $j=1,\ldots,q$, are eigenvectors of G corresponding to σ_i. If for each distinct eigenvalue σ_i we let Y_i denote the set of m_i vectors given by (2.56), then [Halmos (1957), pg. 113] the set of n vectors $\left\{Y_i\right\}_{i=1}^{i=t}$ are linearly independent. Hence, for any matrix G, the set of principal vectors must include a basis for $V_n(C)$.

The integers $q, p_1, p_2, \ldots, p_q$ of Thm. 2.1 may also be given in terms of the elementary divisors of G. Corresponding to the eigenvalue σ_i, the matrix G has the q elementary divisors

$$(z - \sigma_i)^{p_1}, (z - \sigma_i)^{p_2}, \ldots, (z - \sigma_i)^{p_q} . \tag{2.57}$$

Thus, the p_j's, $j=1,\ldots,q$, are simply the degrees of the elementary divisors associated with the eigenvalue σ_i. If all the elementary divisors of G are linear, then the principal vectors of G are also eigenvectors and thus, for this case, the set of eigenvectors includes a basis for $V_n(C)$.

For simplicity reasons, we shall limit ourselves mainly to a discussion of the simplest case which illustrates the character of the changes that occur when the matrix G has nonlinear elementary divisors. Suppose that the matrix G has only one nonlinear elementary divisor and that this nonlinear divisor is of degree 2 and is associated with the eigenvalue σ_s. As a basis for $V_n(C)$, we shall use

$$\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_s, \underline{y}, \underline{x}_{s+1}, \ldots, \underline{x}_{n-1} , \tag{2.58}$$

where the vectors $\underline{x}_i$ are eigenvectors of G with corresponding eigenvalue σ_i and satisfy $G\underline{x}_i = \sigma_i \underline{x}_i$. The vector $\underline{y}$ is a principal vector of grade 2 corresponding to σ_s and satisfies

$$(G - \sigma_s I)\underline{y} = \underline{x}_s \quad (2.59)$$

The existence of such a basis is guaranteed by Theorem 2.1. In what follows we shall assume that the eigenvalues of G are real and that $|\sigma_1| > |\sigma_2| \geq |\sigma_i|$ for $i \geq 3$. We now shall see how the vector $\underline{y}$ affects the convergence rates of the power and Chebyshev polynomial methods of iteration.

B. The Power Method

The eigenvector estimate $\underline{x}(k_1)$ after performing k_1 power iterations may be expressed in terms of the basis vectors (2.58) as

$$\underline{x}(k_1) = \underline{x}_1 + \sum_{i=2} c_i \underline{x}_i + h\underline{y} \, , \quad (2.60)$$

where h and the c_i are scalars. If k_1 is large enough so that the eigenvalue estimates $\sigma(k_1 + r)$, $r \geq 0$, are sufficiently close to σ_1, then for iteration $(k_1 + r)$ we have

$$\underline{x}(k_1 + r) \approx \left(\frac{G}{\sigma_1}\right)^r \underline{x}(k_1) = \underline{x}_1 + \sum_{i=2} \left(\frac{G}{\sigma_1}\right)^r c_i \underline{x}_i + h\left(\frac{G}{\sigma_1}\right)^r \underline{y} \, . \quad (2.61)$$

The corresponding error vector $\underline{E}(k_1 + r) = \underline{x}(k_1 + r) - \underline{x}_1$ may be expressed as

$$\underline{E}(k_1 + r) \approx \sum_{i=2} c_i \left(\frac{G}{\sigma_1}\right)^r \underline{x}_i + h\left(\frac{G}{\sigma_1}\right)^r \underline{y} \, . \quad (2.62)$$

From (2.59), we have that $G\underline{y} = \sigma_s \underline{y} + \underline{x}_s$. Thus,

$$G^2\underline{y} = G[\sigma_s \underline{y} + \underline{x}_s] = (\sigma_s)^2 \underline{y} + 2\sigma_s \underline{x}_s$$

and in general

$$G^r \underline{y} = (\sigma_s)^r \underline{y} + r(\sigma_s)^{r-1} \underline{x}_s \ . \tag{2.63}$$

Hence, $\underline{E}(k_1 + r)$ may be expressed as

$$\underline{E}(k_1 + r) \approx \sum_{i=2} c_i \left(\frac{\sigma_i}{\sigma_1}\right)^r \underline{x}_i + h\left(\frac{\sigma_s}{\sigma_1}\right)^r \underline{y} + \frac{hr}{\sigma_1}\left(\frac{\sigma_s}{\sigma_1}\right)^{r-1} \underline{x}_s \ . \tag{2.64}$$

Since $\lim_{r\to\infty} r\alpha^r = 0$ if $|\alpha| < 1$, we have that $\lim_{r\to\infty} \underline{E}(k_1 + r) = \underline{0}$. Thus, the power method is still a convergent process when principal vectors of grade 2 are present.

The most slowly decaying basis vector in (2.64) will depend on the magnitude of $|\sigma_s/\sigma_1|$ and is likely to vary with r. After r power iterations, the $\underline{x}_s$ vector has a coefficient of

$$c_s r \left(\frac{\sigma_s}{\sigma_1}\right)^{r-1} \left[\frac{\sigma_s}{r\sigma_1} + \frac{h}{\sigma_1 c_s}\right]$$

as compared with a coefficient of c_s at the beginning of these r iterations. Hence, the $\underline{x}_s$ vector has been multiplied by a factor of

$$M_r = r\left(\frac{\sigma_s}{\sigma_1}\right)^{r-1} \left[\frac{\sigma_s}{r\sigma_1} + \frac{h}{\sigma_1 c_s}\right] \tag{2.65}$$

in r power iterations. The multiplication factor for the other basic vectors in the error vector expansion is, as before, $(\sigma_i/\sigma_1)^r$. We note that the vector $\underline{x}_s$ is being built up if the multiplication factor is greater than unity and being reduced when this factor is less than unity.

Although $\lim_{r\to\infty} M_r = 0$, initially M_r may increase with r. For example, if $|\sigma_s/\sigma_1| = .99$ and $h/\sigma_1 c_s = 1.0$, then M_r is an increasing function of r for $r \leq 90$ and is greater than unity for $r \leq 640$. If the ratio $|\sigma_s/\sigma_1|$ is small, however, then M_r goes to zero very rapidly. For example, if $\sigma_s = 0$, then $M_r = 0$ for $r \geq 2$. Thus, for the power method of iteration, the effect of principal vectors of grade 2 depends rather strongly on the value of the corresponding eigenvalue.

C. The Chebyshev Polynomial Method

Suppose that the r-th degree real polynomial given by (2.10) is applied to the eigenvector estimate $\underline{x}(k_1)$. We may write this as

$$\underline{x}(k_1 + r) = P_r[g(G)]\underline{x}(k_1) \quad , \tag{2.66}$$

where $g(z) = \frac{2z}{\sigma_1(d - b)} - \frac{d + b}{d - b}$ and $P_r[g(z)] = \frac{T_r[g(z)]}{T_r[g(\sigma_1)]}$, and where $T_r(g)$ is the Chebyshev polynomial of degree r in g. We note that the notation used in (2.66) is slightly different from that used previously. We have introduced the dependent variable g merely for notational ease later.

With the vector $\underline{x}(k_1)$ given by (2.60), the vector $\underline{x}(k_1 + r)$ can be written as

$$\underline{x}(k_1 + r) = \underline{x}_1 + \sum_{i=2} c_i P_r[g(G)]\underline{x}_i + hP_r[g(G)]\underline{y} \tag{2.67}$$

and the error vector $\underline{E}(k_1 + r)$ as

$$(2.68) \qquad \underline{E}(k_1 + r) = \sum_{i=2} c_i P_r[g(G)]\underline{x}_i + hP_r[g(G)]\underline{y} \ .$$

Since the $\underline{x}_i$ are eigenvectors, the sum $\sum_{i=2} c_i P_r[g(G)]\underline{x}_i$ can be written as $\sum_{i=2} c_i P_r[g(\sigma_i)]\underline{x}_i$. We now want to see what happens to the term $P_r[g(G)]\underline{y}$.

Since $P_r[g(G)]$ is a polynomial of degree r in $g(G)$, we may write

$$(2.69) \qquad P_r[g(G)] = a_0 + a_1[g(G)] + a_2[g(G)]^2 + \cdots + a_r[g(G)]^r \ .$$

Now

$$g(G)\underline{y} = \left[\frac{2G}{\sigma_1(d - b)} - \frac{d + b}{d - b}\right]\underline{y} = g(\sigma_s)\underline{y} + \frac{2}{\sigma_1(d - b)}\,\underline{x}_s \ .$$

Thus,

$$[g(G)]^2\underline{y} = [g(\sigma_s)]^2\underline{y} + 2\left(\frac{2}{\sigma_1(d - b)}\right)g(\sigma_s)\underline{x}_s$$

and in general

$$(2.70) \qquad [g(G)]^r\underline{y} = [g(\sigma_s)]^r\underline{y} + r\left(\frac{2}{\sigma_1(d - b)}\right)[g(\sigma_s)]^{r-1}\underline{x}_s \ .$$

Since $r\left(\frac{2}{\sigma_1(d - b)}\right)[g(\sigma_s)]^{r-1} = \left.\frac{d[g(z)]}{dz}\right|_{z=\sigma_s}$, we have from Eqs. (2.69) and (2.70) that

$$(2.71) \qquad P_r[g(G)]\underline{y} = P_r[g(\sigma_s)]\underline{y} + \frac{d\{P_r[g(\sigma_s)]\}}{dz}\,\underline{x}_s \ .$$

We also have

$$(2.72) \quad \frac{d\{P_r[g(\sigma_s)]\}}{dz} = \frac{dg}{dz}\frac{d[P_r(g)]}{dg} = \left[\frac{2}{\sigma_1(d - b)}\right]\left[\frac{1}{T_r[g(\sigma_1)]}\right]\frac{d[T_r(g)]}{dg} \quad .$$

But [National Bureau of Standards (1952), page ix] $\frac{dT_r(g)}{dt} = rU_{r-1}(g)$, where $U_{r-1}(g)$ is the Chebyshev polynomial of the second kind. $U_{r-1}(g)$ is a polynomial of degree r - 1 and is given by

$$(2.73) \quad U_{r-1}(g) = \frac{\sin(r\cos^{-1}g)}{\sqrt{1 - g^2}} = \binom{r}{1}g^{r-1} - \binom{r}{3}g^{r-3}(1-g^2) + \binom{r}{5}g^{r-5}(1-g^2)^2 + \ldots$$

Thus, using (2.71) we may write (2.68) as

$$(2.74) \quad \underline{E}(k_1+r) = \sum_{i=2} c_i P_r[g(\sigma_i)]\underline{x}_i + hP_r[g(\sigma_s)]\underline{y} + hr\left(\frac{2}{\sigma_1(d-b)}\right)\frac{U_{r-1}[g(\sigma_s)]}{T_r[g(\sigma_1)]}\,\underline{x}_s \quad .$$

Since $b \leq \sigma_s \leq d$, $g(\sigma_s)$ can lie between -1 and +1. Thus, from Eq. (2.73) we see that $|U_{r-1}(g(\sigma_s))| < r$. Since $1/T_r[g(\sigma_1)]$ behaves as α^r, where $|\alpha| < 1$, we have that $\lim_{r\to\infty} \underline{E}(k_1 + r) = \underline{0}$. Thus, the Chebyshev polynomial method is still a convergent process when principal vectors of grade 2 are present in the set of basis vectors.

The most slowly decaying basis vector in (2.74) will depend on the values of σ_s and r. Except for the vector $\underline{x}_s$, all basis vectors in the expansion (2.74) are modified in the normal Chebyshev way. In applying the r-th degree Chebyshev polynomial to $\underline{x}(k_1)$, we see from Eq. (2.74) that the $\underline{x}_s$ vector has been multiplied by a factor of

$$(2.75) \qquad M_r[P_r] = \frac{rU_{r-1}[g(\sigma_s)]}{T_r[g(\sigma_1)]} \left\{ \frac{T_r[g(\sigma_s)]}{rU_{r-1}[g(\sigma_s)]} + \frac{2h}{c_s\sigma_1(d - b)} \right\} ,$$

where we assume that $U_{r-1}[g(\sigma_s)] \neq 0$.

In Table 2.3, we give the values of M_r and $M_r[P_r]$ when $d = .99$, $b = 0$, and $h/c_s\sigma_1 = 1$. In the first case we take $\sigma_s = \sigma_2$ and in the second case we take $\sigma_s = 0$. Note that the magnitude of the multiplication factor of the Chebyshev polynomial method did not change much for the two cases[1] whereas that corresponding to the power method did. Also from Table 2.3, we see that the Chebyshev polynomial iterations would diverge if a polynomial of degree less than 40 were repeatedly applied.

The presence of principal vectors of grade 2 or higher in the set of basis vectors also makes it very difficult to estimate the parameters needed for the efficient use of the Chebyshev polynomial method. In general, it is felt that extreme caution should be exercised when using the Chebyshev polynomial method of iteration if the set of eigenvectors for G does not span the associated vector space.

[1]We remark that $M_r[P_r]$ would be much smaller for the $\sigma_s = 0$ case if we had taken $b = -d$.

r	CASE 1; $\sigma_s = \sigma_2$		CASE 2; $\sigma_s = 0.0$	
	M_r	$M_r[P_r]$	M_r	$M_r[P_r]$
1	1.990	2.960	1.000	1.000
2	2.960	8.396	.000	-6.547
3	3.911	16.163	.000	14.478
4	4.842	24.871	.000	-23.378
5	5.754	35.878		34.585
10	10.040	53.615		-53.087
20	17.341	29.233		-29.161
40	27.698	1.277		- 1.276
80	36.611	.003		- .003
160	32.370			
320	12.965			
640	1.040			

TABLE 2.3

D. Principal Vectors of Grade Greater Than Two

Suppose now that the nonlinear elementary divisor of G is of degree $m + 1$, where m is arbitrary. For this case we take as a basis for $V_n(C)$, the n vectors

$$\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_s, \underline{y}_1, \underline{y}_2, \ldots, \underline{y}_m, \underline{x}_{s+1}, \ldots, \underline{x}_t \;, \tag{2.76}$$

where the vectors $\underline{x}_i$ are eigenvectors of G with corresponding eigenvalue σ_i and the vector $\underline{y}_j$, $j=1,2,\ldots,m$ is a principal vector of grade $j + 1$ corresponding to the eigenvalue σ_s. The vectors $\underline{y}_j$ satisfy the relationship $\underline{y}_j = (G-\sigma_s I)^{m-j}\underline{y}_m$ or equivalently

$$(2.77)\qquad \begin{cases} (G - \sigma_s I)\underline{y}_m = \underline{y}_{m-1} \\ (G - \sigma_s I)\underline{y}_{m-1} = \underline{y}_{m-2} \\ \vdots \\ (G - \sigma_s I)\underline{y}_2 = \underline{y}_1 \\ (G - \sigma_s I)\underline{y}_1 = \underline{x}_s \end{cases}$$

The existence of such a basis is guaranteed by Theorem 2.1. We also assume that the eigenvalues of G are real and that $|\sigma_1| > |\sigma_2| \geq |\sigma_i|$ for $i \geq 3$.

In terms of the basis vectors (2.76), the eigenvector estimate after k_1 iterations may be expressed as

$$(2.78)\qquad \underline{x}(k_1) = \underline{x}_1 + \sum_{i=2} c_i\underline{x}_i + \sum_{j=1}^{m} h_j\underline{y}_j \quad ,$$

where we again assume that the eigenvalue estimates $\sigma(k_1 + r)$, $N \geq 0$, are sufficiently close to σ_1.

For the power method of iteration, the error vector $\underline{E}(k_1 + r)$ may be expressed as

$$\underline{E}(k_1 + r) \approx \sum_{i=2}\left(\frac{G}{\sigma_1}\right)^r c_i\underline{x}_i + \sum_{j=1}^{m}\left(\frac{G}{\sigma_1}\right)^r h_j\underline{y}_j$$

which after some manipulation may be written as

$$(2.79) \qquad \underline{E}(k_1 + r) \approx \sum_{i=2} \left(\frac{\sigma_i}{\sigma_1}\right)^r c_i \underline{x}_i + \left(\frac{\sigma_s}{\sigma_1}\right)^{r-m} \left[D_0 \underline{x}_s + \sum_{j=1}^{m} D_j \underline{y}_j \right] ,$$

where

$$D_j = \frac{1}{(\sigma_1)^m} \sum_{k=0}^{m-j} \binom{r}{k} (\sigma_s)^{m-k} h_{j+k} \quad , \quad j=0,1,\ldots,m \ ,$$

and where h_0 is taken to be zero. Since σ_1, σ_s, m, and the h_j's are independent of r and finite, there exists nonnegative constants $\tilde{D}_j$ such that

$$(2.80) \qquad |D_j| \leq \tilde{D}_j \binom{r}{j} \quad , \quad j=0,1,\ldots,m \quad .$$

Thus, since $\lim_{r \to \infty} (r)^m (\sigma_s/\sigma_1)^{r-m} = 0$, again we have that $\lim_{r \to \infty} \underline{E}(k_1 + r) = \underline{0}$.

From (2.79) and (2.80), we see that the coefficient of the $\underline{x}_s$ basis vector in the expansion (2.78) goes to zero as

$$(2.81) \qquad \binom{r}{m} \left(\frac{\sigma_s}{\sigma_1}\right)^{r-m} \quad .$$

Thus, for large m, the presence of principal vectors of grade m may greatly reduce the convergence rate of the power method.

If the eigenvector estimate $\underline{x}(k_1 + r)$ had been obtained by applying the r-th degree polynomial $P_r[g(G)]$ of Eq. (2.66) to $\underline{x}(k_1)$, then the error vector $\underline{E}(k_1 + r)$ may be expressed as

$$(2.82) \qquad \underline{E}(k_1 + r) \approx \sum_{i=2} c_i P_r[g(G)] \underline{x}_i + \sum_{j=1}^{m} h_j P_r[g(G)] \underline{y}_j \ .$$

In a manner similar to that given previously for the case m = 1, we obtain

$$\underline{E}(k_1 + r) \approx \sum_{i=2} c_i P_r[g(\sigma_i)]\underline{x}_i + \sum_{j=1}^{m} h_j P_r[g(\sigma_s)]\underline{y}_j$$

(2.83) $$+\left\{\sum_{j=1}^{m} \frac{h_j}{j!} \frac{d^j\{P_r[g(\sigma_s)]\}}{dz^j}\right\}\underline{x}_s + \left\{\sum_{j=1}^{m-1} \frac{h_{j+1}}{j!} \frac{d^j\{P_r[g(\sigma_s)]\}}{dz^j}\right\}\underline{y}_1$$

$$+ \cdots + h_m \frac{d\{P_r[g(\sigma_s)]\}}{dz} \underline{y}_{m-1} .$$

To show that the error vector given by (2.83) approaches the null vector as r approaches infinity, one needs to show that

(2.84) $$\lim_{r \to \infty} \frac{d^j\{P_r[g(\sigma_s)]\}}{dz^j} = 0$$

for $j=0,1,\ldots,m$. With $g(\sigma_1) \geq 2.0$ or equivalently with $d \leq \frac{2}{3} + \frac{b}{3}$, one may easily show that (2.84) is true. With $d < 1.0$, we conjecture that (2.84) is true.

We now turn to the practical problems of estimating the constants d and b, and of terminating the iterative process.

III. THE ESTIMATION OF THE DOMINANCE RATIO d AND THE TERMINATION OF THE ITERATIVE PROCESS

In this chapter we shall give a criterion for terminating the iterative process (2.12) and shall specify a numerical means by which to estimate the dominance ratio d. For the purposes of this chapter we shall assume that the eigenvalues of the n×n matrix G are real and are ordered such that

$$\sigma_n \leq \sigma_{n-1} \leq \cdots \leq \sigma_3 \leq \sigma_2 \leq \sigma_1$$

and that the set of eigenvectors for G includes a basis for $V_n(C)$. We also assume that $\sigma_1 > \sigma_2 > |\sigma_i|$ for $i \geq 3$.

1. The Estimation of d and b

As mentioned previously, the use of non-optimum values for d and b can result in a sizable reduction in the convergence rate of the Chebyshev method of iteration. Thus, one is faced with the problem of determining these unknown constants in order to use the Chebyshev polynomial method efficiently.

Since we have assumed that $d > |b|$, the rate of convergence of the Chebyshev polynomial method will be governed primarily by the value of d. Thus, the estimate for d appears to be the more important estimate. Hence, in what follows we assume that enough is known about the eigenvalues of G so that an estimate b_0 for b may be picked to satisfy $b_0 < b$ and $|b_0| < d$. We remark that these conditions on the choice of b_0 are not impractical. For if the eigenvalues of G are all positive, then $b > 0$ and $b_0 = 0$ is a satisfactory choice. If nothing is known about b, then one may apply the Chebyshev polynomial method with the argument matrix G^2 instead of G. Since the eigenvalues of G^2 must be nonnegative, zero is a lower bound for the eigenvalues of G^2. We remark that the effective convergence rate of the Chebyshev polynomial method with argument matrix G^2 differs very little from that with argument matrix G. In fact, if $b = -d$, then the effective convergence rates of the two are identical. This follows from the identity $T_{2r}(x) = T_r(2x^2 - 1)$.

The rate of convergence will not be critically affected by the estimate b_0 if b_0 satisfies the above two conditions. This then is why we feel justified in assuming that b_0 satisfy only the two conditions given above and that an accurate estimate for b is not essential.[1] Thus, henceforth, we will be concerned only with estimates for d.

In order to obtain an accurate estimate for d, we propose the following strategy. Before starting the Chebyshev method of iteration, do a few (say 5 or 10) power iterations in order to obtain an initial estimate for d. (These initial power iterations also provide a reasonable estimate for σ_1 for use in the initial Chebyshev iterations.) Then apply repeatedly low degree Chebyshev polynomials so that the estimates for d may be continuously updated. After a good estimate for d is obtained high degree polynomials may be applied, if needed.

Numerical estimates for d may be obtained by observing the decay rate of the residual vector $\underline{y}(k) = \underline{v}(k) - \underline{x}(k - 1)$, where $\underline{v}(k)$ and $\underline{x}(k - 1)$ are defined by (2.12). We define the residual vector quotient as

---

[1]It is essential, though, that a $b_0 \leq b$ be used. For if $b_0 > b + (1 - d)$, then the Chebyshev polynomial method will diverge.

$$(3.1) \qquad Q(k) = \frac{\|\underline{y}(k)\|}{\|\underline{y}(k-1)\|} \; ,$$

where $\| \; \|$ denotes some suitable vector norm. For the power method, it is known that $\lim_{k\to\infty} Q(k) = d$. To see why this is true, let $\underline{x}(0)$ be expanded in terms of the eigenvectors of G as

$$\underline{x}(0) = \underline{x}_1 + \sum_{i=2}^{n} c_i \underline{x}_i \; .$$

Thus, we have

$$\underline{v}(k) - \underline{x}(k-1) = \frac{(\sigma_1)^{k-1}}{\sigma(0)\cdots\sigma(k-2)} \left\{ \left(\frac{\sigma_1}{\sigma(k-1)} - 1\right)\underline{x}_1 + \left(\frac{\sigma_2}{\sigma(k-1)} - 1\right) d^{k-1} c_2 \underline{x}_2 \right.$$

$$\left. + \sum_{i=3}^{n} \left(\frac{\sigma_i}{\sigma(k-1)} - 1\right)\left(\frac{\sigma_i}{\sigma_1}\right)^{k-1} c_i \underline{x}_i \right\} \; .$$

Since $\lim_{k\to\infty} \sigma(k) = \sigma_1$ and $|\sigma_i|/\sigma_1 < d$ for $i \geq 3$, we have that $\lim_{k\to\infty} Q(k) = d$. Thus, an initial estimate for d may be obtained by doing a few power iterations before starting the use of Chebyshev polynomials.

Estimates for d may also be obtained every Chebyshev iteration by comparing the convergence rate actually being obtained with the theoretical convergence rate one would obtain if the d being used were correct. If a Chebyshev polynomial using d_0 as the estimate for d is started on iteration $k_1 + 1$ and if

$$(3.2)\quad \begin{cases} \text{(i)} & \sigma(k_1 + r) \text{ is sufficiently close to } \sigma_1 \text{ for all } r \geq 0 \\ & \text{and} \\ \text{(ii)} & \left\|\sum_{i=3}^{n} P_{r,d_0}(\sigma_i/\sigma_1)c_i\underline{x}_i\right\| \text{ is small relative to } \left\|P_{r,d_0}(d)c_2\underline{x}_2\right\|, \end{cases}$$

then from Eq. (2.9) we may approximate $\underline{x}(k_1 + r)$ by

$$\underline{x}(k_1 + r) \approx P_{r,d_0}(1)\left\{\underline{x}_1 + \frac{P_{r,d_0}(d)}{P_{r,d_0}(1)}\, c_2\underline{x}_2\right\},$$

where $P_{r,d_0}(y)$ is given by Eq. (2.17). Since $P_{r,d_0}(1) = 1$, the residual vector $\underline{y}(k_1 + r + 1)$ may be approximated by $\underline{y}(k_1 + r + 1) \approx (d - 1)P_{r,d_0}(d)c_2\underline{x}_2$ and the residual vector quotient by

$$(3.3)\qquad Q(k_1 + r + 1) \approx \left|\frac{P_{r,d_0}(d)}{P_{r-1,d_0}(d)}\right| ,$$

where $P_{0,d_0}(y) = 1$. With $Q_{r+1} \equiv \prod_{\ell=2}^{r+1} Q(k_1 + \ell)$, it follows that

$$(3.4)\qquad Q_{r+1} \approx |P_{r,d_0}(d)| .$$

Thus, Q_{r+1} gives a measure of $|P_{r,d_0}(d)|$ and one may obtain a new estimate for the dominance ratio by solving (3.4) for d. We now shall describe how one may obtain a new estimate for d from (3.4).

<u>Case 1</u>: $\qquad 1 > Q_{r+1} > P_{r,d_0}(d_0)$

From Figure 3.1, this case implies that $d > d_0$ and we are not obtaining the expected convergence rate from the present estimate d_0. Thus, a new estimate for d should be obtained for possible use in the generation of a new Chebyshev polynomial. To obtain this new estimate for d we make use of (3.4). Using $P_{r,d_0}(d)$ as defined in (2.17), we may express (3.4) as

$$(3.5) \qquad Q_{r+1} \approx \frac{T_r\left(\dfrac{2d - d_0 - b_0}{d_0 - b_0}\right)}{T_r\left(\dfrac{2 - d_0 - b_0}{d_0 - b_0}\right)}$$

or equivalently since $P_{r,d_0}(d_0) = 1.0/T_r\left(\dfrac{2 - d_0 - b_0}{d_0 - b_0}\right)$

$$(3.6) \qquad T_r\left(\frac{2d - d_0 - b_0}{d_0 - b_0}\right) \approx \frac{Q_{r+1}}{P_{r,d_0}(d_0)} \quad .$$

The right side of (3.6) is greater than one so that the largest positive solution to (3.6) can be expressed as

$$(3.7) \qquad d \approx \left(\frac{d_0 - b_0}{2}\right)\left\{\cosh\left[\frac{\cosh^{-1}\left\{\dfrac{Q_{r+1}}{P_{r,d_0}(d_0)}\right\}}{r}\right] + \left(\frac{d_0 + b_0}{d_0 - b_0}\right)\right\} .$$

This solution may then be used as the new estimate for d. One may easily show that the d given by (3.7) satisfies the inequality $d_0 < d < 1$.

<u>Case 2</u>: $\qquad Q_{r+1} < P_{r,d_0}(d_0)$

From Figure 3.1, we see that this case implies that $d < d_0$ and we are getting a convergence rate which is greater than that expected from using d_0. For this case the right side of (3.6) is less than one and a solution to (3.6) is

(3.8) $$d \approx \left(\frac{d_0 - b_0}{2}\right)\left\{\cos\left[\frac{\cos^{-1}\left\{\frac{Q_{r+1}}{P_{r,d_0}(d_0)}\right\}}{r}\right] + \left(\frac{d_0 + b_0}{d_0 - b_0}\right)\right\}.$$

If the principal value is used for the inverse cosine, then the resulting d estimate satisfies $\mu \leq d < d_0$, where μ is the largest positive root of $P_{r,d_0}(y) = 0$. (See Figure 3.1.)

Case 3: $Q_{r+1} > 1.0$

This case implies that there has been no error reduction. If the error is not being reduced, then one or more of our assumptions are not being satisfied. This can happen

(a) if $\sigma(k_1 + r)$ is not a sufficiently good approximation of σ_1,

(b) if the set of eigenvectors of G does not span $V_n(C)$,

(c) If the eigenvalues of G are not all real, or

(d) if $b < b_0$.

If the assumptions given in (3.2) are valid, expressions (3.7) and (3.8) normally will give a good estimate for d. Obviously, these assumptions do not always hold. However, they may be reasonable under certain conditions. The Chebyshev strategy given below is designed toward this end.

2. Chebyshev Strategy

Basically, the Chebyshev strategy can be divided into three parts, as follows:

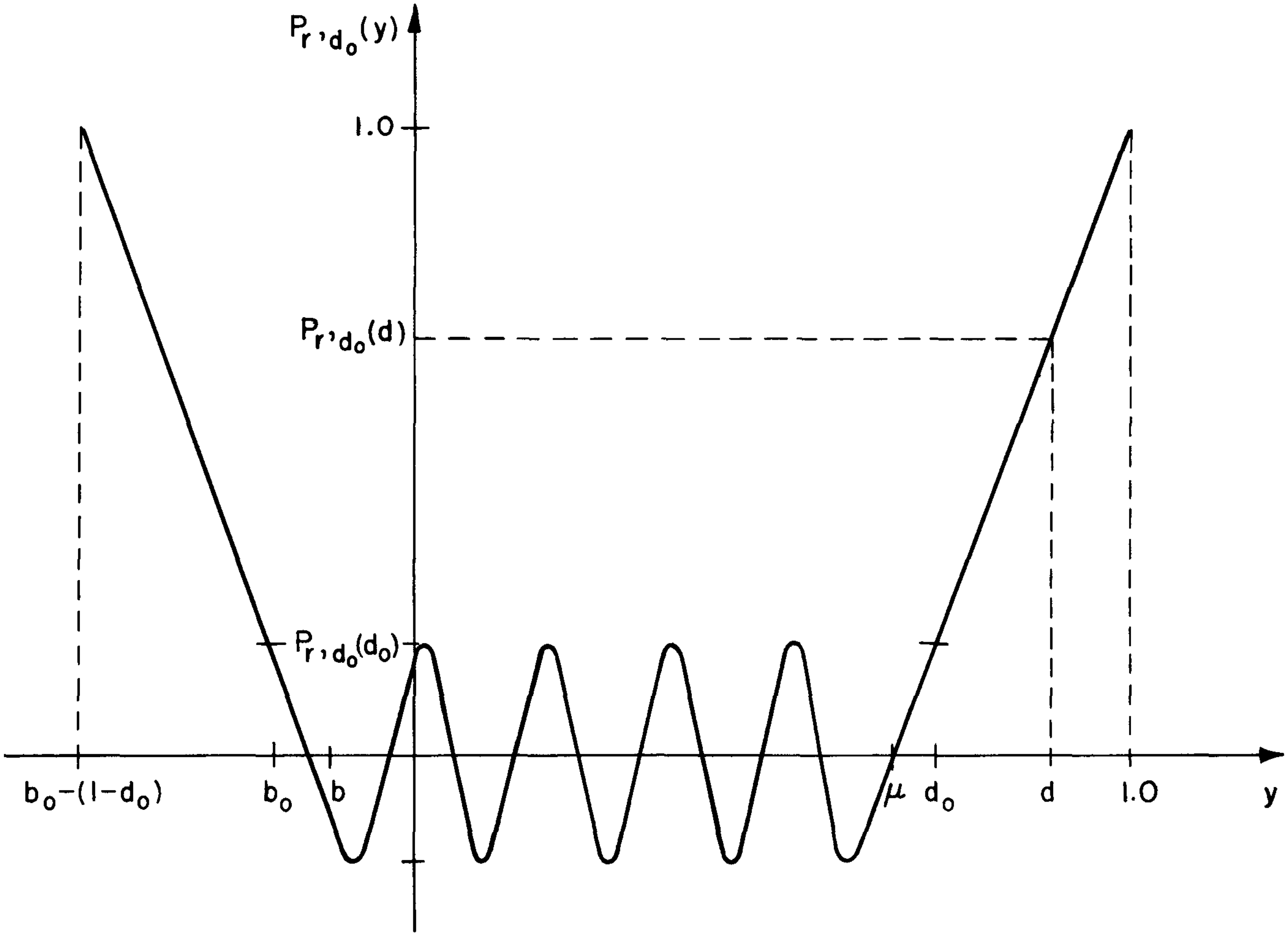


FIG. 3.1

(a) Initially, at least four iterations of the power type are carried out in order to obtain an initial estimate d_0 for d and a reasonable estimate for σ_1. We note that these power iterations will practically eliminate from the eigenvector guess $\underline{x}(0)$ those eigenvector modes corresponding to the smaller eigenvalues. (See Eq. (2.5)).

(b) The use of Chebyshev polynomials is then started on iteration 5, say, using d_0 as the estimate for d. Low degree Chebyshev polynomials are repeatedly applied with the estimates for the dominance ratio being continuously updated. If the low degree Chebyshev polynomials are generated with the dominance ratio under-estimated, these polynomials will greatly reduce all the eigenvector modes in the guess vector $\underline{x}(0)$ except those with the larger eigenvalues. For example, if $d = .889$ and if a 5-th degree Chebyshev polynomial is generated with $d_0 = .8$, then all eigenvector modes $\underline{x}_i$ with $(\sigma_i/\sigma_1) \leq .8$ are multiplied by a factor smaller in magnitude than .017, while the $\underline{x}_2$ eigenvector mode is multiplied by a factor of only .211. Thus, generating a polynomial with the dominance ratio under-estimated results in assumption (ii) of (3.2) being more nearly satisfied. One may impose upper bounds on the initial d estimates in an effort to make these estimates less than d. For example, one may insist that $d_0 \leq .9$, $d_1 \leq .925$, etc.

(c) As (3.2) becomes more nearly satisfied giving relatively good convergence towards the correct d, high degree Chebyshev polynomials may be applied, if needed, to reduce those eigenvector modes with the larger eigenvalues.

In summary, the Chebyshev strategy is to first eliminate the more rapidly decaying eigenvector modes from the guess $\underline{x}(0)$ and then concentrate on the most slowly decaying modes. This generally enables the estimates for d to converge to the correct value.

The decision whether to terminate the present Chebyshev polynomial and start the generation of a new polynomial using an improved estimate for d can be made by comparing the convergence rate actually being obtained with the theoretical convergence rate one would obtain if the estimate for d were correct.

The convergence rate for iteration k is defined to be

$$\theta(k) \equiv -\ln \frac{\|\underline{E}(k)\|}{\|\underline{E}(k-1)\|} \tag{3.9}$$

where $\underline{E}(k) \equiv \underline{x}(k) - \underline{x}_1$ is the error vector for iteration k. If a Chebyshev polynomial using d_0 as the estimate for d is started on iteration k+1, then from Eq. (2.9) we may write $\underline{E}(k + r) = \sum_{i=2} c_i P_{r,d_0}(\sigma_i/\sigma_1)\underline{x}_i$. Using assumption (3.2) and Eq. (3.3), $\theta(k + r)$ may be approximated by

$$\theta(k + r) \approx -\ln\left|\frac{P_{r,d_0}(d)}{P_{r-1,d_0}(d)}\right| \approx -\ln[Q(k + r + 1)] .$$

Now if $d_0 = d$, the Chebyshev theory of Chapter II implies that the theoretical convergence rate for iteration k + r should be $-\ln[P_r(d)/P_{r-1}(d)]$ or equivalently $-\ln[T_{r-1}(a)/T_r(a)]$, where $a = (2 - d_0 - b_0)/(d_0 - b_0)$. Thus,

$$R(k + r + 1) = \frac{\ln[Q(k + r + 1)]}{\ln[T_{r-1}(a)/T_r(a)]} \tag{3.10}$$

may be used to compare the actual convergence rate with the theoretical convergence rate for iteration k + r.

The decision whether to begin a new Chebyshev polynomial using a new estimate for d can be based on R(k). For example, one could start a new polynomial on iteration $k + r + 1$ if $R(k + r + 1)$ is less than .7. It is often helpful to insist that all polynomials generated be at least of degree r^*, where r^* may be taken to be 3 or 4.

3. Terminating the Iterative Procedure

Let the relative sum error $\Lambda(k)$ for iteration k be defined as

$$\Lambda(k) \equiv \frac{\|\underline{E}(k)\|_2}{\|\underline{x}_1\|_2} \quad , \tag{3.11}$$

where by $\|\underline{r}\|_2$ is meant the Euclidean or ℓ_2 norm of the vector r, i.e., $\|\underline{r}\|_2 \equiv [\underline{r},\underline{r}]^{1/2}$.

In order to obtain a computable approximation for $\Lambda(k)$ let

$$\Delta(k) \equiv \frac{\|\underline{v}(k) - \underline{x}(k - 1)\|_2}{\|\underline{x}(k - 1)\|_2} \quad . \tag{3.12}$$

We now assume k is large enough so that $\underline{x}(k - 1)$ may be approximated by

$$\underline{x}(k - 1) \approx \underline{x}_1 + c_2\underline{x}_2 \quad . \tag{3.13}$$

Thus $\Delta(k)$ may be expressed as

$$\Delta(k) \approx \left\{ \frac{(1 - \delta)^2 c_2^2 \underline{x}_2,\underline{x}_2}{\underline{x}_1,\underline{x}_1 + 2c_2\underline{x}_1,\underline{x}_2 + c_2^2\underline{x}_2,\underline{x}_2} \right\}^{1/2} \quad . \tag{3.14}$$

Using (3.13), we have $\underline{E}(k - 1) = \underline{x}(k - 1) - \underline{x}_1 \approx c_2\underline{x}_2$ so that $\Lambda(k - 1) \approx |c_2| \cdot \|\underline{x}_2\|/\|\underline{x}_1\|$ and

$$\Delta(k) \approx \Lambda(k - 1) \frac{(1 - d)}{\{1 + y + [\Lambda(k - 1)]^2\}^{1/2}} ,$$

where $y = 2c_2[\underline{x}_1,\underline{x}_2]/[\underline{x}_1,\underline{x}_1]$. Using Schwarz's inequality we have $|y| \leq 2\Lambda(k - 1)$. Thus, if $\Lambda(k - 1) < 1$, we have

$$(3.15) \qquad [1 - \Lambda(k - 1)] \lesssim \frac{(1 - d)\Lambda(k - 1)}{\Delta(k)} \lesssim [1 + \Lambda(k - 1)] .$$

Hence, for k sufficiently large we have

$$(3.16) \qquad \frac{\Delta(k)}{(1 - d) + \Delta(k)} \lesssim \Lambda(k - 1) \lesssim \frac{\Delta(k)}{(1 - d) - \Delta(k)} .$$

Another possible measure as to how well $\underline{x}(k)$ approximates $\underline{x}_1$ is what we shall call the <u>relative point error</u> $\lambda(k)$. If $\underline{e}_j$ is a vector of order n whose j-th component is unity and all other components zero, then the relative point error for iteration k is defined as

$$(3.17) \qquad \lambda(k) \equiv \max_j \left| \frac{[\underline{e}_j,\underline{E}(k)]}{[\underline{e}_j,\underline{x}_1]} \right| ,$$

where the subscript j varies only over the set of indices for which $(\underline{e}_j,\underline{x}_1) \neq 0$.

To obtain a computable approximation for $\lambda(k)$, we let

$$(3.18) \qquad \delta(k) \equiv \max_j \left| 1 - \frac{[\underline{e}_j,\underline{v}(k)]}{[\underline{e}_j,\underline{x}(k-1)]} \right| ,$$

where the subscript j again varies only over the set of indices for which

$(\underline{e}_j, \underline{x}(k-1)) \neq 0.$[1]

Again assuming that k is sufficiently large and using Eq. (3.13), we have

$$\frac{[\underline{e}_j, \underline{v}(k)]}{[\underline{e}_j, \underline{x}(k-1)]} \approx \frac{(\underline{e}_j, \underline{x}_1) + (d)c_2(\underline{e}_j, \underline{x}_2)}{(\underline{e}_j, \underline{x}_1) + c_2(\underline{e}_j, \underline{x}_2)}$$

and if $(\underline{e}_j, \underline{x}_1) \neq 0$, then

$$\frac{[\underline{e}_j, \underline{v}(k)]}{[\underline{e}_j, \underline{x}(k-1)]} \approx 1 - \frac{(1-d)a_j}{1+a_j} , \tag{3.19}$$

where

$$a_j = c_2 \frac{(\underline{e}_j, \underline{x}_2)}{(\underline{e}_j, \underline{x}_1)} .$$

Thus, $\delta(k)$ may be approximated by

$$\delta(k) \approx (1-d) \max_j \left| \frac{a_j}{1+a_j} \right| . \tag{3.20}$$

Using (3.13), $\lambda(k-1)$ may be approximated by $\lambda(k-1) \approx \max_j |a_j|$. Thus, if $\max_j |a_j| < 1$, we have

$$\frac{\lambda(k-1)}{1+\lambda(k-1)} \lesssim \frac{\delta(k)}{1-d} \lesssim \frac{\lambda(k-1)}{1-\lambda(k-1)} \tag{3.21}$$

---

[1] In practice, one usually may avoid those j for which $(\underline{e}_j, \underline{x}_1) = 0$ by allowing the indices j in (3.18) to vary only over those j for which $[\underline{e}_j, \underline{x}(k-1) \geq \gamma\{\max_j[\underline{e}_j, \underline{x}(k-1)]\}$, where γ is some fixed small number.

and

$$(3.22) \qquad \frac{\delta(k)}{(1-d)+\delta(k)} \lesssim \lambda(k-1) \lesssim \frac{\delta(k)}{(1-d)-\delta(k)}$$

Thus, one could terminate the iterative procedure (2.12) by using $\Delta(k)$ and/or $\delta(k)$ modified in some way by a function of d to measure the relative sum and point errors. We note that the relative sum error is an aggregate measure of the error vector $\underline{E}(k)$ while the relative point error is a pointwise measure.

$\Delta(k)$ may also be used to estimate the relative eigenvalue error $\tau(k)$, where

$$(3.23) \qquad \tau(k) \equiv \left| \frac{\sigma(k)}{\sigma_1} - 1 \right| .$$

From Eq. (2.12), $\sigma(k+1)/\sigma_1$ is given by

$$\frac{\sigma(k+1)}{\sigma_1} = \frac{\sigma(k)}{\sigma_1} \frac{[\underline{v}(k+1),\underline{v}(k+1)]}{[\underline{v}(k+1),\underline{x}(k)]} = \frac{[\frac{G}{\sigma_1}\underline{x}(k),\frac{G}{\sigma_1}\underline{x}(k)]}{[\frac{G}{\sigma_1}\underline{x}(k),\underline{x}(k)]}$$

and hence

$$(3.24) \qquad \frac{\sigma(k+1)}{\sigma_1} - 1 = \frac{[\frac{G}{\sigma_1}\underline{x}(k),(\frac{G}{\sigma_1}\underline{x}(k) - \underline{x}(k))]}{[\frac{G}{\sigma_1}\underline{x}(k),\underline{x}(k)]} .$$

If k is large enough so that the vector estimate $\underline{x}(k)$ can be written as $\underline{x}(k) \approx \underline{x}_1 + c_2\underline{x}_2$ then (3.24) can be approximated by

$$(3.25) \qquad \frac{\sigma(k+1)}{\sigma_1} - 1 \approx \frac{c_2(d-1)[\underline{x}_1,\underline{x}_2 + c_2 d\underline{x}_2,\underline{x}_2]}{\underline{x}_1,\underline{x}_1 + c_2(1-d)\underline{x}_1,\underline{x}_2 + c_2^2 d\underline{x}_2,\underline{x}_2} .$$

But $[\Lambda(k)]^2 \approx c_2^2[\underline{x}_2,\underline{x}_2]/[\underline{x}_1,\underline{x}_1]$ so that

$$\left|\frac{\sigma(k+1)}{\sigma_1} - 1\right| \lesssim (1-d)\,\Lambda(k)\left[\frac{1 + d\Lambda(k)}{(1 - d\Lambda(k)(1-\Lambda(k))}\right]$$

and using (3.15) and the fact that $d < 1$, we have

$$\left|\frac{\sigma(k+1)}{\sigma_1} - 1\right| \lesssim \left[\frac{1+\Lambda(k)}{1-\Lambda(k)}\right]^2 \Delta(k+1) \quad . \tag{3.26}$$

The inequalities (3.16), (3.22), and (3.26) are based on the assumption that k is large enough so that

1. the eigenvalue estimates $\sigma(k)$ are sufficiently close to σ_1 and that
2. the eigenvector expansion of the error vector $\underline{E}(k)$ consists of one predominant eigenvector.

The conditions given above were needed in order to give some mathematical basis for these inequalities. It is felt, however, that the indicated bounds are realistic under much less stringent conditions. In using (3.16) and (3.22), it is important that one have a good estimate for d. This is especially true when d is close to unity.

We note that the inequality (3.26) may be sharpened somewhat if the matrix G is symmetric. For this case the set of basis vectors may be chosen to be orthogonal. Thus, since $\underline{x}_1,\underline{x}_2 = 0$, Eqs. (3.14) and (3.25) may be expressed as

$$\Delta(k+1) \approx \left\{\frac{(1-d)^2 c_2^2 \underline{x}_2,\underline{x}_2}{\underline{x}_1,\underline{x}_1 + c_2^2\underline{x}_2,\underline{x}_2}\right\}^{1/2} = \frac{(1-d)\,\Lambda(k)}{[1+(\Lambda(k))^2]^{1/2}}$$

and

$$\frac{\sigma(k - 1)}{\sigma_1} - 1 \approx \frac{c_2^2 d(d - 1)\underline{x}_2,\underline{x}_2}{\underline{x}_1,\underline{x}_1 + c_2^2 d\underline{x}_2,\underline{x}_2} \approx \frac{d(d - 1)[\Lambda(k)]^2}{1 + d[\Lambda(k)]^2} \text{ .}$$

Hence, since $d < 1$, we get

$$\left|\frac{\sigma(k + 1)}{\sigma_1} - 1\right| \lesssim \frac{[\Delta(k + 1)]^2}{1 - d} \text{ .} \tag{3.28}$$

In the next chapter, we shall discuss some numerical results.

IV. NUMERICAL EXAMPLES

In this chapter we give numerical examples which illustrate certain points concerning the behavior of the Chebyshev polynomial method.

We seek to solve the homogeneous problem

$$Gx = \sigma x \tag{4.1}$$

for the dominant eigenvalue σ_1 and its corresponding eigenvector $\underline{x}_1$. The Chebyshev iterations are carried out using the procedure

$$(4.2)\quad \begin{cases} \underline{v}(k_1 + t) = \dfrac{G}{\sigma(k_1 + t - 1)} \underline{x}(k_1 + t - 1) \\[2ex] \sigma(k_1 + t) = \sigma(k_1 + t - 1) \dfrac{[\underline{v}(k_1 + t), \underline{v}(k_1 + t)]}{[\underline{v}(k_1 + t), \underline{x}(k_1 + t - 1)]} \\[2ex] x(k_1 + t) = \underline{x}(k_1 + t - 1) + \alpha_{k_1+t}\left[\dfrac{\sigma(k_1 + t - 1)}{\sigma(k_1 + t)} \underline{v}(k_1 + t) - \underline{x}(k_1 + t - 1)\right] \\[2ex] \qquad\qquad + \beta_{k_1+t}\left[\underline{x}(k_1 + t - 1) - \underline{x}(k_1 + t - 2)\right], \end{cases}$$

where α_{k_1+t} and β_{k_1+t} are given by (2.13). The above procedure differs from (2.12) only in the normalization of [illegible] for $\underline{v}(k_1 + t)$ in the extrapolation. For convergent problems, numerical experiments indicate that both procedures (2.12) and (4.2) give essentially the same results.

The Chebyshev strategy is basically that as described in Chapter III. The generation of a Chebyshev polynomial is terminated and a new polynomial started on iteration k if

$$(4.3)\quad \begin{cases} \text{(i)} & \text{the degree of the polynomial to be terminated is} \\ & \text{greater than or equal to 3, and if} \\ \text{(ii)} & R(k) < .6 \ , \end{cases}$$

where $R(k)$ is given by (3.10). If it is decided to terminate the n-th polynomial then d_{n+1}, the estimate for d to be used in the generation of the (n + 1) Chebyshev polynomial, is usually taken to be the d as determined from (3.7) or (3.8). However, the following restrictions are placed on d_{n+1}: $d_0 < .95$, $d_1 \leq .985$, $d_2 \leq .995$, and $d_n \leq .99995$ for $n \geq 3$.

In the numerical data given below, we let

k = the iteration index.

$\sigma(k)$ = the estimate for σ_1 after k iterations.

$d(k)$ = the estimate for d after k iterations.

r = the degree of the Chebyshev polynomial which has been generated at the end of the k-th iteration. r = 0 implies that the k-th iteration is a power iterate.

$R(k)$ = the ratio of the actual convergence rate to the theoretical convergence rate for iteration k - 1. $R(k) \equiv 1.0$ if iteration k - 1 was a power iteration; otherwise $R(k)$ is as defined by (3.10).

$\Delta(k)$ = the error estimate as defined by (3.12).

$\delta(k)$ = the error estimate as defined by (3.18).

For problem 1, the matrix G of (4.1) is a symmetric, positive semi-definite matrix of order 99 whose nonzero eigenvalues are $\sigma_\ell = [\cos \pi\ell/100]^2$, $\ell=1,2,\ldots,49$. The remaining 50 eigenvalues of G are equal to zero. Thus, $\sigma_1 = .999013$ and $d = \sigma_2/\sigma_1 = .99704$. Since the matrix G is symmetric for this problem, the set of eigenvectors for G includes an orthonormal basis for the associated vector space. Thus, the real Chebyshev polynomial method of Chapter II may be rigorously applied.

Three different iteration strategies were used to obtain the dominant eigenvalue for problem 1. First the problem was run using the Chebyshev strategy described above. Then the dominant eigenvalue was obtained using only power iterations and finally the problem was solved using only one high degree Chebyshev polynomial with the correct value for d being input. The numerical results for problem 1 are given in Tables 4.1-4.3. Graphs of $\Delta(k)$ vs. k for the three iteration strategies are given in Figure 4.1.

Comments on Problem 1:

From Table 4.1, we see that the Chebyshev strategy worked as expected for this problem.

The odd behavior of $R(k)$, $\Delta(k)$, and $\delta(k)$ in Table 4.3 around the 25th and 50th iterations is probably due to the relationship between the eigenvalues σ_ℓ, of G and the zeros and/or peaks of the Chebyshev polynomials of degree 24 and 49. Generally, if a high degree Chebyshev polynomial using an input value for d which is too large is generated early, the behavior of the R, Δ, δ and d quantities is much more erratic. This is illustrated in problem 3 below.

PROBLEM 1

CHEBYSHEV ACCELERATION

k	r	R(k)	α(k)	Δ(k)	δ(k)	d(k)
2	0	1.00	.989613	.04652	.24965	.35350
5	1	1.00	.993801	.01917	.09994	.82201
7	3	.41	.995651	.01066	.03517	.89142
8	1	.23	.996157	.00872	.02702	.90706
9	2	.55	.996404	.00787	.02740	.94661
10	3	.41	.996810	.00651	.02602	.95272
11	1	.28	.997113	.00553	.02047	.95726
13	3	.60	.997455	.00453	.01367	.97211
14	1	.47	.997679	.00388	.01092	.97434
16	3	.68	.997869	.00338	.01086	.98144
17	1	.59	.998021	.00297	.00927	.98258
20	4	.69	.998247	.00240	.00625	.98687
21	1	.60	.998349	.00214	.00559	.98758
25	5	.65	.998543	.00166	.00418	.99097
26	1	.58	.998604	.00151	.00362	.99132
32	7	.65	.998784	.00103	.00226	.99345
33	1	.60	.998814	.00093	.00197	.99360
41	9	.64	.998921	.00057	.00126	.99501
42	1	.59	.998932	.00052	.00111	.99511
50	9	.64	.998971	.00036	.00073	.99638
51	1	.60	.998976	.00034	.00065	.99643
60	10	.80	.998997	.00022	.00048	.99696
61	11	.78	.998999	.00021	.00046	.99697
62	12	.76	.999001	.00019	.00045	.99698
90	40	.60	.999013	.00002	.00005	.99704

TABLE 4.1

PROBLEM 1

POWER ITERATIONS

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	0	1.00	.989613	.04652	.24965	.35350
5	0	1.00	.993801	.01917	.09994	.82201
7	0	1.00	.994808	.01442	.07129	.87426
8	0	1.00	.995155	.01292	.06236	.89593
9	0	1.00	.995439	.01173	.05542	.90854
10	0	1.00	.995678	.01077	.04987	.91841
11	0	1.00	.995883	.00998	.04532	.92634
13	0	1.00	.996215	.00873	03833	.93831
14	0	1.00	.996353	.00823	.03558	.94294
16	0	1.00	.996588	.00741	.03112	.95036
17	0	1.00	.996689	.00706	.02928	.95339
20	0	1.00	.996945	.00621	.02486	.96057
21	0	1.00	.997017	.00598	.02367	.96249
25	0	1.00	.997260	.00522	.01986	.96858
26	0	1.00	.997312	.00506	.01909	.96980
32	0	1.00	.997567	.00430	.01547	.97546
33	0	1.00	.997602	.00420	.01500	.97620
41	0	1.00	.997836	.00354	.01203	.98077
42	0	1.00	.997860	.00347	.01174	.98121
50	0	1.00	.998026	.00303	.00983	.98411
51	0	1.00	.998040	.00298	.00963	.98441
60	0	1.00	.998182	.00262	.00815	.98661
61	0	1.00	.998195	.00258	.00802	.98681
90	0	1.00	.998473	.00187	.00535	.99061
300	0	1.00	.998935	.00053	.00129	.99566

TABLE 4.2

PROBLEM 1

CHEBYSHEV (HIGH DEGREE) ACCELERATION - INPUT d = .99704

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	1	1.00	.989613	.04652	.24965	.99704
5	4	1.08	.995101	.02976	1.19120	.98015
7	6	.98	.996114	.02488	.25943	.99096
8	7	.99	.996627	.02231	.16182	.99283
9	8	.99	.997082	.01985	.11047	.99395
10	9	.99	.997464	.01759	.08013	.99468
11	10	.99	.997776	.015551	.06005	.99517
23	22	1.02	.998908	.00371	.01066	.99670
24	23	1.03	.998921	.00330	.00959	.99672
25	24	1.04	.998929	.00294	.00861	.99675
26	25	-.42	.998940	.00306	.00931	.99688
27	26	4.30	.998962	.00192	.00653	.99667
28	27	-1.57	.998968	.00226	.00674	.99685
29	28	1.50	.998973	.00192	.00617	.99684
30	29	.99	.998979	.00172	.00567	.99685
50	49	.58	.999012	.00022	.01807	.99700
51	50	8.44	.999013	.00009	00746	.99695
52	51	-7.86	.999013	.00020	.02578	.99703
53	52	3.99	.999013	.00013	.00577	.99697
54	53	.19	.999013	.00013	.00816	.99700
55	54	.99	.999013	.00012	.00239	.99700
56	55	.98	.999013	.00010	.00141	.99700
57	56	.98	.999013	.00009	.00105	.99700
58	57	.99	.999013	.00008	.00081	.99700
71	70	.96	.999013	.00002	.00008	99701

TABLE 4.3

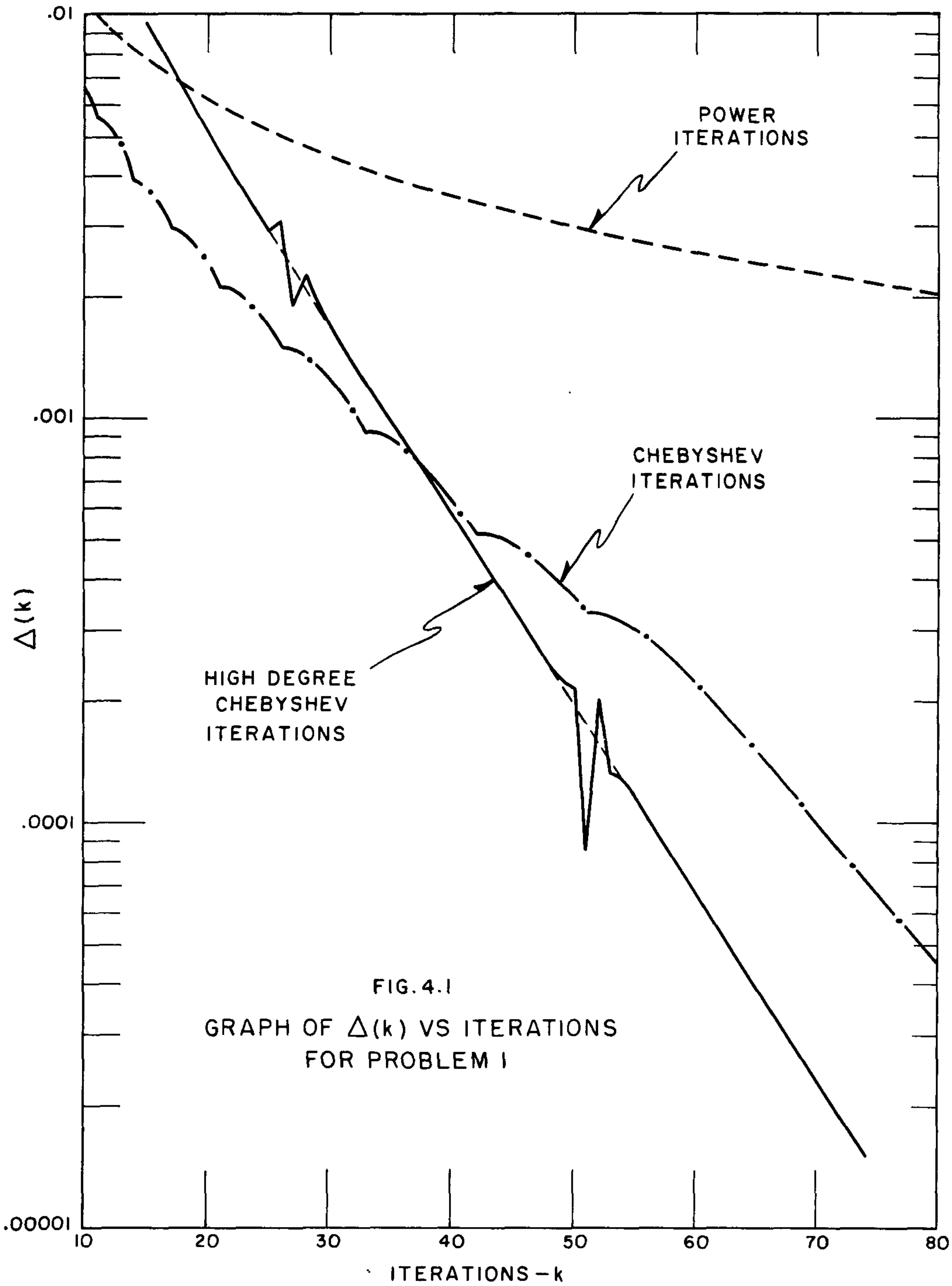


FIG. 4.1

GRAPH OF Δ(k) VS ITERATIONS FOR PROBLEM 1

For problem 2, the matrix G has the same eigenvalues as those for problem 1 but the set of eigenvectors for this problem does not include a basis for the associated vector space. A nonlinear elementary divisor of degree 50 is associated with the eigenvalue zero and thus, principal vectors of grade 50 are present in the set of basis vectors for G. The numerical results obtained from this problem are given in Tables 4.4-4.7. In Tables 4.6 and 4.7, the quantity Q(k) is defined by Eq. (3.1).

Comments on Problem 2:

The power iterations for problem 2 were almost identical to those of problem 1. This would not have been the case, however, had the nonlinear divisor been associated with some rather large nonzero eigenvalue.

Note the rapid divergence property of the Chebyshev iterations. The results of the Chebyshev high degree acceleration problems do not necessarily imply that the conjecture of (2.84) is false. In (2.84) it is assumed that an infinite degree Chebyshev polynomial is generated and that the true dominant eigenvalue is used as the normalizing factor in the Chebyshev extrapolation. In an attempt to see the effect of the normalizing factor, the high degree Chebyshev acceleration problem of Table 4.6 was rerun using the true dominant eigenvalue in the Chebyshev extrapolation, i.e., the $\underline{x}(k_1 + t)$ term in (4.2) was computed using $[\sigma(k_1 + t - 1)/\sigma_1]\underline{v}(k_1 + t)$ instead of $[\sigma(k_1 + t - 1)/\sigma(k_1 + t)]\underline{v}(k_1 + t)$. As seen from Tables 4.6 and 4.7, this change seemed to affect only the eigenvalue estimates. The reason for the relative stability of the R(k), d(k), and Q(k) quantities in Tables 4.6 and 4.7 is not known.

The numerical results of problem 2 show that the presence of nonlinear elementary divisors can drastically affect the behavior of the Chebyshev polynomial method. We remark, however, that the Chebyshev polynomial method often may be used to good advantage when the nonlinear divisors are associated with very small eigenvalues and when the strategy used makes provisions for the presence of principal vectors in the set of basis vectors. For a discussion of this, see Hageman and Kellogg (1966).

PROBLEM 2

CHEBYSHEV ACCELERATION

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	0	1.00	.988735	.04155	.24904	.38662
5	0	1.00	.993636	.01712	.09973	.82724
7	3	.43	.995671	.00952	.03573	.89279
8	1	.22	.996184	.00785	.02622	.90927
9	2	.49	.996411	.00718	.02684	.95319
10	3	.34	.996807	.00615	.02957	.96015
11	1	.20	.997118	.00548	.03706	.96575
13	3	-1.42	.997475	.00715	.08527	
14	1	-2.22	.997743	.01312	.19625	.96575
16	3	-4.23	.998267	.04166	.57306	
17	1	-2.94	.999980	.09300	4.15835	.96575
18	2	-5.94	1.002620	.14019	6.34790	
19	3	-4.32	1.022880	.32197	7.57360	
20	1	-2.92	1.161700	.77061	8.96510	.96575
22	3	-2.38	-7.201190	1.97715	.89157	
23	1	- .19	4.877060	1.68904	1.38180	.96575
26	4	1.23	1.192005	.83736	4.28377	
27	1	-2.03	15.030100	1.94680	1.14667	.96575
51	3	-1.25	7.245400	2.28150	1.09660	

TABLE 4.4

PROBLEM 2

POWER ITERATIONS

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	0	1.00	.988735	.04155	.24904	.38662
5	0	1.00	.993636	.01712	.09973	.82724
7	0	1.00	.994710	.01301	.07123	.88353
8	0	1.00	.995074	.01170	.06232	.89970
9	0	1.00	.995371	.01067	.05539	.91187
10	0	1.00	.995620	.00983	.04984	.92137
11	0	1.00	.995831	.00913	.04530	.92900
13	0	1.00	.996174	.00803	.03832	.94048
14	0	1.00	.996315	.007584	.03557	.94492
50	0	1.00	.998015	.00287	.00983	.98443
51	0	1.00	.998033	.00282	.00963	.98471
90	0	1.00	.998465	.00179	.00535	.99068
300	0	1.00	.998929	.00049	.00128	.99543

TABLE 4.5

PROBLEM 2

CHEBYSHEV (HIGH DEGREE) ACCELERATION INPUT d = .99704

k	r	R(k)	σ(k)	Q(k)	δ(k)	d(k)
3	1	1.00	.991	.6488	.1662	.99704
5	3	-.08	.996	1.0015	.4545	.94707
7	5	19.10	.998	2.1296	1.5511	
8	6	18.45	1.0001	2.4899	1.1931	
9	7	16.33	1.018	2.5046	1.9539	
10	8	14.43	1.180	2.6049	1.3407	
11	9	11.93	-1.098	2.3974	1.9913	
25	23	- .12	8.149	1.0125	1.0044	
26	24	.006	37.401	.9994	1.0155	
27	25	.09	87.462	.9896	1.2289	
51	49	.73	54.195	.9240	1.0416	
52	50	.73	43.764	.9232	1.0005	
53	51	.74	55.134	.9225	1.0007	
100	98	.89	32.889	.9079	1.0091	
101	99	.89	92.769	.9077	1.0014	
102	100	.89	47.791	.9076	1.0033	
200	198	.95	3.920	.9018	1.0001	.9982
201	199	.95	21.413	.9018	1.0030	.9982
202	200	.95	79.812	.9017	1.0007	.9982
298	296	.97	20.268	.9001	1.0022	.9976
299	297	.97	7.289	.9000	1.0006	.9976
300	298	.97	67.839	.9000	1.0002	.9976

TABLE 4.6

PROBLEM 2

CHEBYSHEV (HIGH DEGREE) ACCELERATION - INPUT d = .99704
INPUT σ_1 = .999013

k	r	R(k)	σ(k)	Q(k)	δ(k)	d(k)
3	1	1.00	.991	.6421	.1662	.99704
5	3	-.93	.996	1.0165	.5151	.94677
7	5	19.06	.998	2.1268	1.4248	
8	6	18.44	1.000	2.4895	1.1609	
9	7	16.33	1.021	2.5941	1.7265	
10	8	14.41	1.219	2.6002	1.3235	
11	9	11.68	-.789	2.3534	3.8515	
25	23	.10	57.687	.9898	1.0360	
26	24	.18	43.993	.9810	1.0015	
27	25	.24	4.244	.9743	1.0265	
51	49	.74	42.202	.9226	1.0015	
52	50	.75	53.643	.9219	1.0177	
53	51	.75	23.402	.9212	1.0418	
100	98	.89	52.409	.9077	1.0049	
101	99	.89	11.303	.9076	1.0027	
102	100	.89	60.198	.9074	1.0027	
200	198	.95	97.918	.9018	1.0017	.99805
201	199	.95	85.801	.9018	1.0123	.99804
202	200	.95	76.197	.9017	1.0023	.99803
298	296	.97	33.731	.9001	1.0027	.99752
299	297	.97	50.082	.9000	1.0002	.99752
300	298	.97	28.078	.9000	1.0008	.99752

TABLE 4.7

For problem 3, the matrix G is not symmetric but the set of eigenvectors for G is known to include a basis for the associated vector space. It is also known that the eigenvalues of G are nonnegative. The dominant eigenvalue, σ_1, for this problem appears to be .999886. The numerical results are given in Tables 4.8-4.10 and graphs of $\Delta(k)$ vs. k for the three iteration strategies are given in Figure 4.2.

Comments on Problem 3:

Using Table 4.9, the inequality (3.22) for the relative point error after 300 power iterations gives

$$.365 \lesssim \lambda(299) \lesssim 1.26 \quad .$$

One component of the eigenvector estimate after 300 power iterations had a relative point error of .55. Thus, the inequality (3.22) can give realistic bounds for $\lambda(k)$. In using (3.22), it is well to keep in mind that a sufficiently good estimate for d is needed. For example, using d(k) as the estimate for d, inequality (3.22) gives for iteration 73

$$.055 < \frac{.00118}{.0201 + .00118} \lesssim \lambda(72) < \frac{.00118}{.0201 - .00118} < .063 \ ,$$

which obviously is not correct.

In Table 4.8, note that a Chebyshev polynomial of degree 12 with $d_0 = .97689$ was used early and after that low degree polynomials were again used. This implies that the initial guess vector had a rather large error component associated with an eigenvalue σ_i, where $\sigma_i/\sigma_1 \approx .98$. This fact is also implied from the power iterations since from Table 4.9 we see that the estimate for d from iteration 30 to iteration 73 varied between only .978 and .980. One of the nice properties of the Chebyshev polynomial method is that the method, if properly used, can pick out large components in the error vector and reduce them efficiently.

The high degree Chebyshev problem with input d = .99976 is converging at a slower rate than the Chebyshev problem with the strategy. This is due to the fact that the error components associated with smaller eigenvalues are being reduced in the high degree problem at a rate dictated by the second largest eigenvalue σ_2. The erratic behavior of $\Delta(k)$ for the high degree Chebyshev problem can be easily seen in Figure 4.2.

PROBLEM 3

CHEBYSHEV ACCELERATION

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
5	1	1.00	.997093	.002691	.00493	.90000
7	3	.32	.997686	.002080	.00366	.95686
8	1	.22	.997973	.001820	.00345	.95000
10	3	.40	.998333	.001539	.00313	.97550
11	1	.35	.998594	.001351	.00283	.97689
22	12	.65	.999693	.000168	.00032	.97979
23	1	.53	.999708	.000142	.00025	.98006
25	3	.38	.999718	.000133	.00025	.99110
26	1	.33	.999727	.000126	.00024	.99194
29	4	.65	.999740	.000114	.00022	.99425
30	1	.55	.999748	.000108	.00021	.99470
32	3	.48	.999751	.000105	.00021	.99707
35	3	.59	.999755	.000102	.00021	.99836
38	3	.61	.999759	.000100	.00020	.99904
39	1	.58	.999760	.000099	.00020	.99908
42	4	.61	.999762	.000098	.00020	.99941
43	1	.55	.999763	.000097	.00020	.99945
45	3	.49	.999764	.000097	.00020	.99971
46	1	.46	.999764	.000097	.00019	.99973
52	7	.60	.999768	.000095	.00019	.99982
53	1	.56	.999769	.000095	.00019	.99983
63	8	1.61	.999772	.000094	.00018	.99991
73	18	2.54	.999788	.000088	.00015	.99988
83	28	3.37	.999813	.000076	.00011	.99985
93	38	5.43	.999842	.000055	.00007	.99982
103	48	9.76	.999867	.000028	.00003	.99979
110	55	26.46	.999881	.000009	.00001	.99976

TABLE 4.8

PROBLEM 3

POWER ITERATIONS

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	0	1.00	.996539	.00406	.00899	.44934
5	0	1.00	.997093	.00269	.00493	.92902
7	0	1.00	.997323	.00242	.00377	.95161
8	0	1.00	.997421	.00231	.00374	.95665
23	0	1.00	.998383	.00146	.00303	.97595
30	0	1.00	.998653	.00125	.00269	.97820
43	0	1.00	.999001	.00095	.00213	.97946
44	0	1.00	.999022	.00093	.00209	.97949
45	0	1.00	.999042	.00091	.00205	.97951
52	0	1.00	.999166	.00079	.00179	.97959
53	0	1.00	.999182	.00077	.00176	.97959
63	0	1.00	.999316	.00063	.00144	.97967
73	0	1.00	.999416	.00051	.00118	.97990
83	0	1.00	.999492	.00042	.00096	.98036
93	0	1.00	.999551	.00034	.00078	.98108
103	0	1.00	.999597	.00029	.00064	.98209
110	0	1.00	.999623	.00025	.00056	.98298
150	0	1.00	.999709	.00015	.00026	.99003
200	0	1.00	.999748	.00011	.00021	.99691
250	0	1.00	.999760	.00010	.00020	.99910
300	0	1.00	.999766	.00010	.00019	.99967

TABLE 4.9

PROBLEM 3

CHEBYSHEV (HIGH DEGREE) ACCELERATION - INPUT $\bar{d}$ = .99976

k	r	R(k)	σ(k)	Δ(k)	δ(k)	d(k)
2	1	1.00	.996539	.004064	.00899	.99976
5	4	17.82	.998659	.002883	.00736	.98318
7	6	-1.13	1.000413	.002736	.00620	.99315
8	7	64.65	1.001129	.003823	.01215	.99932
10	9	7.06	1.000933	.002878	.00718	.99767
11	10	25.62	1.001089	.003512	.01233	.99926
12	11	11.67	1.000705	.003150	.00963	.99893
13	12	-4.55	.999991	.003277	.00942	.99927
14	13	20.41	.999961	.004045	.01264	
15	14	36.02	.999909	.002604	.00532	.99891
42	41	-6.49	.999502	.001475	00353	99971
43	42	-5.55	.999638	.001698	.00461	.99977
44	43	-6.84	.999799	.002027	.00611	.99983
45	44	9.16	.999772	.001578	.00484	.99976
46	45	11.45	.999651	.001153	.00287	.99970
53	52	.02	1.000180	.001130	.00443	.99975
63	62	3.49	.999403	.000697	.00162	.99973
73	72	10.13	.999860	.000564	.00147	.99975
83	82	-2.53	.999802	.000485	.00135	.99976
93	92	-2.41	1.000034	.000299	.00072	.99975
103	102	16.17	.999890	.000273	.00091	.99976
107	106	.92	.999862	.000198	00044	.99976
108	107	22.47	.999830	.000099	.00031	.99973
109	108	-8.46	.999808	.000128	.00032	99974
110	109	-7.93	.999819	.000164	.00038	.99975
111	110	1.98	.999842	.000154	.00034	.99975

TABLE 4.10

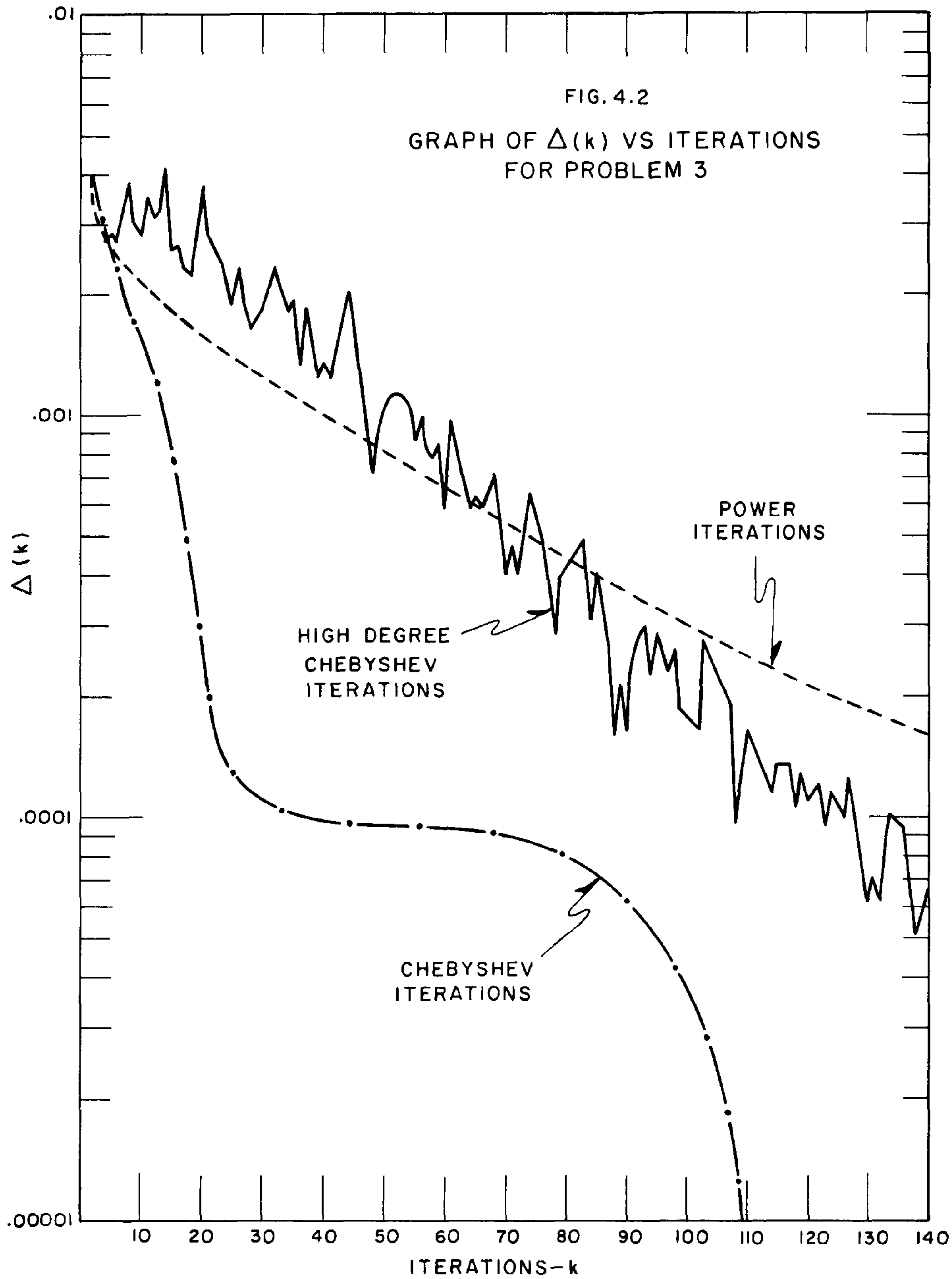


FIG. 4.2
GRAPH OF Δ(k) VS ITERATIONS
FOR PROBLEM 3
POWER
ITERATIONS
HIGH DEGREE
CHEBYSHEV
ITERATIONS
CHEBYSHEV
ITERATIONS
.01
.001
.0001
.00001
Δ (k)
10
20
30
40
50
60
70
80
90
100
110
120
130
140
ITERATIONS—k

For problem 4, the only fact known concerning the matrix G is that it has a positive dominant eigenvalue. The purpose of this problem is to illustrate the behavior of the complex domain Chebyshev polynomial method described in Chapter II. The numerical results for this problem were obtained from a program designed for a different purpose and, hence, the strategy employed is slightly different and fewer numbers are available as output.

The complex domain Chebyshev method requires the use of the three parameters b, d, and ϵ of Figure 2.4. For this problem, b was assumed equal to -d. The problem was run three times using complex Chebyshev acceleration: first with fixed $\epsilon = .1$, then with fixed $\epsilon = .5$, and finally with fixed $\epsilon = .75$. In Tables 4.11-4.14, r and d(k) are defined as before. However, $\delta(k)$ is now defined as

$$\delta(k) \equiv \max_j \left| 1 - \frac{[\underline{e}_j, \underline{x}(k)]}{[\underline{e}_j, \underline{x}(k-2)]} \right|$$

and Q(k) as

$$Q(k) \equiv \frac{\| \underline{y}(k) \|}{\| \underline{y}(k-2) \|}$$

where $\underline{y}(k) = \underline{v}(k) - \underline{x}(k-2)$. The $\delta(k)$ and Q(k) defined above have the same basic meanings as given previously, i.e., $\delta(k)$ is still a measure of the relative point error and Q(k) can be used to estimate d. For power iterations we have $\lim_{k \to} Q(k) = d^2$. In the Chebyshev accelerated problems, the estimate for the dominance ratio is not updated at each iteration.

Graphs of $\delta(k)$ vs. k are given in Figures 4.3 and 4.4.

Comments on Problem 4:

The Q(k) quantity in Table 4.11 indicates that the matrix G for this problem has rather large complex eigenvalues. To see why this is so, we shall assume that G has some large complex eigenvalues and show that the behavior of Q(k) is the same as that of Table 4.11.

Let the eigenvalues of G be ordered such that $\sigma_1 > \sigma_2 > |\sigma_3| = |\sigma_4| > |\sigma_5| \geq \cdots$, where $\sigma_4 = \bar{\sigma}_3$, i.e., we are assuming σ_3 and σ_4 to be complex eigenvalues. Also, let $\underline{x}_i$ be the corresponding eigenvectors. Now if k is sufficiently large so that $\underline{x}(k - 2)$ can be approximated by $\underline{x}(k - 2) \approx \underline{x}_1 + \underline{x}_2 + \underline{x}_3 + \bar{\underline{x}}_3$, then $\underline{y}(k + m)$ can be approximated by

$$\underline{y}(k + m) \approx d^m(d^2 - 1)\underline{x}_2 + z^m(z^2 - 1)\underline{x}_3 + \overline{z^m(z^2 - 1)\underline{x}_3} , \tag{4.4}$$

where $z = \sigma_3/\sigma_1$. If we denote the j-th component of the vector $(z^2 - 1)\underline{x}_3$ by $a_j e^{i\beta_j}$ and the j-th component of the vector $(d^2 - 1)\underline{x}_2$ by x_j, then the j-th component of $\underline{y}(k + m)$ can be written as

$$y_j(m) \approx d^m[x_j + 2r^m a_j \cos(m\theta + \beta_j)] , \tag{4.5}$$

where $z \equiv |z|e^{i\theta}$ and $r \equiv |z|/d$. Thus, we have

$$||\underline{y}(k + m)||^2 \equiv \sum_j [y_j(m)]^2 \approx d^{2m} \sum_j [x_j^2 + r^m b_j \cos(m\theta + \beta_j)] ,$$

where $b_j = 4a_j x_j$ and where we have assumed the r^{2m} term in $[y_j(m)]^2$ to be negligible. Thus, we have

$$[Q(k+m)]^2 \approx d^4 \left[\frac{1 + r^m \sum_j c_j \cos(m\Theta + \beta_j)}{1 + r^{m-2} \sum_j c_j \cos((m-2)\Theta + \beta_j)} \right]$$

where $c_j = b_j \Big/ \sum_j x_j^2$. Again neglecting terms of order greater than r^m, we have

$$[Q(k+m)]^2 \approx d^4 \left[1 + r^{m-2} \sum_j c_j [r^2(\cos(m\Theta + \beta_j) - \cos((m-2))\Theta + \beta_j)] \right] .$$

which may be written as

$$[Q(k+m)]^2 \approx d^4 \left\{ 1 + r^{m-2} [s(\cos m\Theta) + t(\sin m\Theta)] \right\} ,$$

where $s = \sum_j c_j [r^2 \cos \beta_j - \cos(\beta_j - 2\Theta)]$ and where $t = \sum_j c_j [\sin(\beta_j - 2\Theta) - r^2 \sin \beta_j]$. If $s^2 + t^2 \neq 0$, then $s(\cos m\Theta) + t(\sin \Theta) = v \sin(m\Theta + \phi)$, where $v = (s^2 + t^2)^{1/2}$ and $\Theta = \sin^{-1} s/c$. Thus, we have

$$(4.6) \qquad [Q(k+m)]^2 \approx d^4[1 + vr^{m-2} \sin(m\Theta + \phi)] .$$

Thus, $Q(k + m)$ will oscillate about d^2 with a period of $2\pi/\Theta$. Since we have assumed $r < 1$, the amplitude of this oscillation becomes smaller with m. This is precisely the behavior of $Q(k)$ in Table 4.11. The peaks in $Q(k)$ have been underlined in Table 4.11

From Table 4.11, we see that the period is between 38 and 42 so that the argument Θ of $z(=\sigma_3/\sigma_1)$ is roughly between 8.5° and 9.5°. The numerical data of Table 4.11 also implies that the modulus of z is about .93. The Chebyshev data given in Tables 4.13 and 4.14 imply that the normalized complex eigenvalues of G lie interior to the ellipse

$$\frac{x^2}{(.946)^2} + \frac{y^2}{(.75)^2} = 1 \tag{4.7}$$

but exterior to

$$\frac{x^2}{(.968)^2} + \frac{y^2}{(.5)^2} = 1 \; . \tag{4.8}$$

The complex number $z = .93e^{i(2\pi/40)}$ is interior to both ellipses but $z = .935e^{i(2\pi/38)}$ is interior to the ellipse of (4.7) and exterior to the ellipse of (4.8).

PROBLEM 4

POWER ITERATIONS

k	r	δ(k)	Q(k)
12	0	.8312430	.470567
32	0	.0334180	1.187410
52	0	.0097594	.783629
72	0	.0035820	.997618
112	0	.0004183	.937051
152	0	.0000494	.914422
154	0	.0000435	.914765
156	0	.0000384	.913669
158	0	.0000339	.911347
160	0	.0000301	.908208
162	0	.0000269	.904545
164	0	.0000242	.900857
166	0	.0000218	.897353
168	0	.0000198	.894374
170	0	.0000180	.892083
172	0	.0000164	.890686
174	0	.0000150	.890124
176	0	.0000136	.890562
178	0	.0000124	.891785
180	0	.0000112	.893693
182	0	.0000102	.895988
184	0	.0000092	.898412
186	0	.0000082	.900863
188	0	.0000074	.902887
190	0	.0000066	.904489
192	0	.0000059	.905593
194	0	.0000052	.905985
196	0	.0000047	.905442

TABLE 4.11

PROBLEM 4

COMPLEX DOMAIN CHEBYSHEV ACCELERATION - FIXED $\epsilon = .1$

k	r	δ(k)	Q(k)	d(k)
10	0	.954786	.590	.768
18	8	.442435	.787	
20	0	.258927	.593	.913
22	2	.146929	.887	
38	18	.175314	.674	
40	0	.112929	1.394	.900
42	2	.149251	.913	
44	4	.148386	1.005	
46	6	.286167	2.155	
48	8	.485156	1.300	
50	10	.971887	.843	
52	12	1.617660	1.986	
54	14	4.354200	.713	
56	16	3.90364	1.587	
58	18	84.09410	1.317	
60	20	54.00010	.893	
62	22	22.94770	1.539	
64	24	27.97570	1.139	
66	26	18.38360	.602	
68	28	85.22920	1.727	
70	0	16.95400	.903	.900
72	2	75.16020	.833	
100	30	53.60490	2.095	

TABLE 4.12

PROBLEM 4

COMPLEX DOMAIN CHEBYSHEV ACCELERATION - FIXED $\epsilon = .5$

k	r	δ(k)	Q(k)	d(k)
10	0	.954786	.590	.768
18	8	.427459	.809	
20	0	.290486	.738	.906
22	2	.207921	.867	
38	18	.022285	.541	
40	0	.018444	.347	.967
42	2	.019464	.549	
44	4	.025491	1.405	
46	6	.029178	1.116	
48	8	.028033	.890	
50	10	.017541	.479	
52	12	.012640	1.328	
54	14	.011752	1.601	
56	16	.006478	1.104	
58	18	.012880	1.221	
60	20	.009910	.972	
62	22	.008962	.404	
64	24	.007952	1.084	
66	26	.008748	1.592	
68	28	.006889	.309	
70	0	.007033	2.764	.900
72	2	.004470	.886	
108	38	.001312	2.097	
110	0	.000825	.594	.968
112	2	.000701	.768	
158	48	.006667	.854	
198	38	.001256	.451	

TABLE 4.13

PROBLEM 4

COMPLEX DOMAIN CHEBYSHEV ACCELERATION - FIXED $\epsilon = .75$

k	r	δ(k)	Q(k)	d(k)
10	0	.9547860	.590	.768
18	8	.4150640	.827	
20	0	.3155640	.844	.899
22	2	.2434230	.864	
38	18	.0231445	.973	
40	0	.0193948	.823	.900
42	2	.0156790	.840	
44	4	.0154278	.845	
46	6	.0128705	.731	
48	8	.0118865	.730	
50	10	.0120340	.741	
52	12	.0128080	.768	
54	14	.0127994	.803	
56	16	.0119710	.841	
58	18	.0103050	.885	
60	20	.0082678	.934	
62	22	.0063066	.984	
64	24	.0046881	1.019	
66	26	.0034826	1.030	
68	28	.0026561	1.014	
70	0	.0020136	.911	.941
72	2	.0019455	.947	
108	38	.0002086	.870	
110	0	.0001729	.766	.946
112	2	.0001714	.865	
158	48	.0000110	.886	
160	0	.0000085	.803	.946

TABLE 4.14

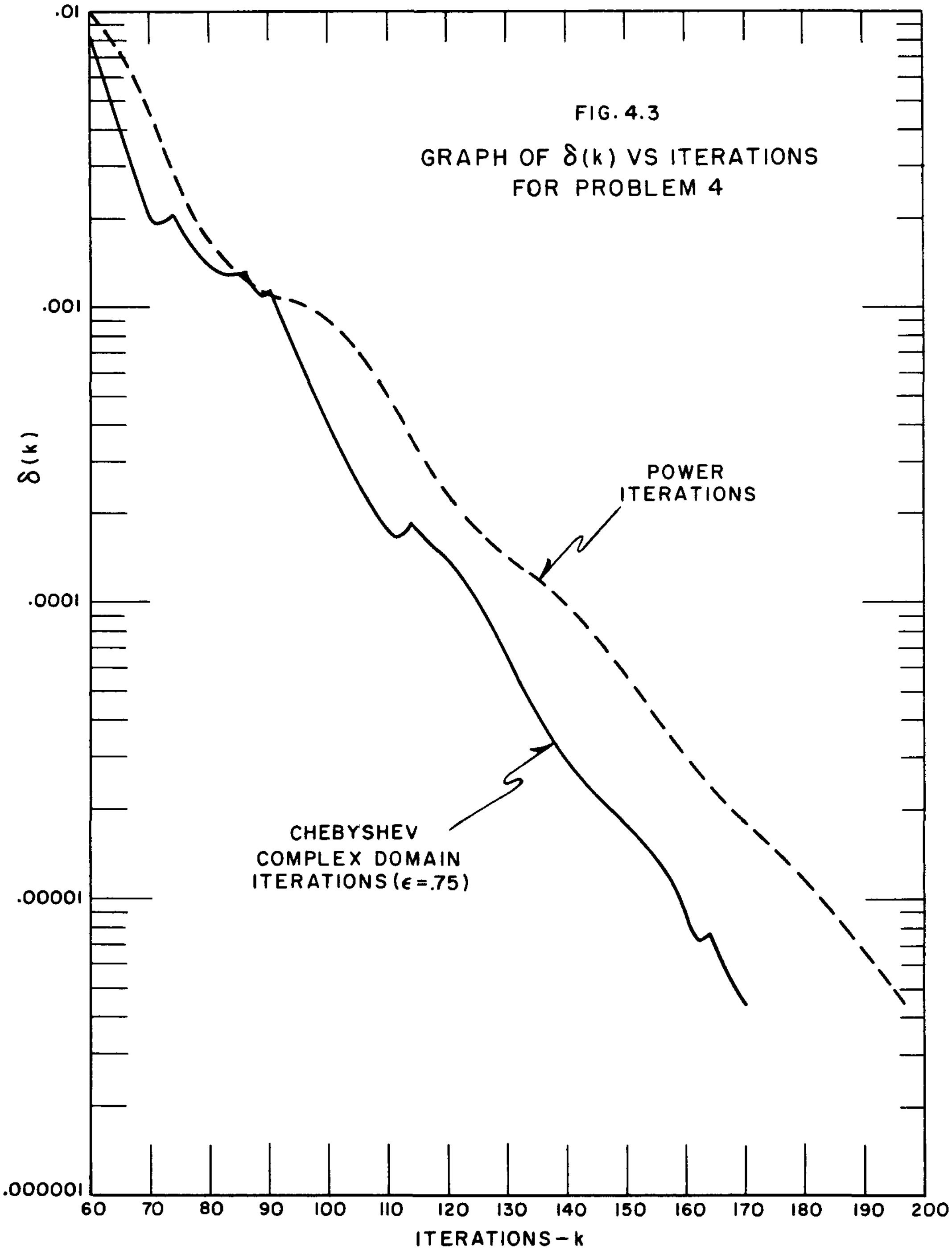


FIG. 4.3

GRAPH OF δ(k) VS ITERATIONS FOR PROBLEM 4

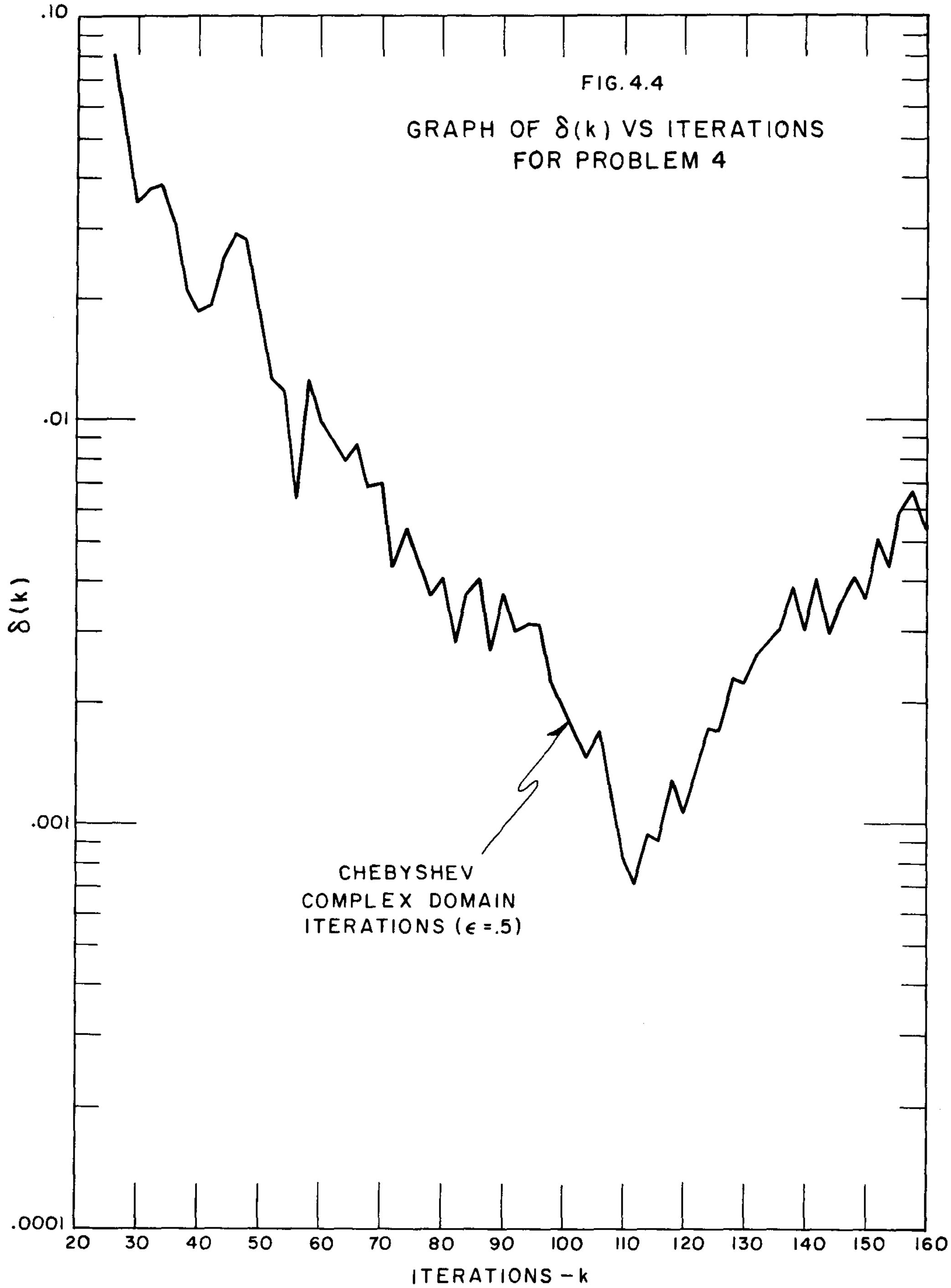


FIG. 4.4
GRAPH OF δ(k) VS ITERATIONS
FOR PROBLEM 4
.10
.01
.001
.0001
δ(k)
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
ITERATIONS – k
CHEBYSHEV
COMPLEX DOMAIN
ITERATIONS (ε =.5)

APPENDIX A

THE REAL DOMAIN CHEBYSHEV POLYNOMIAL AND COMPLEX EIGENVALUES

In this appendix, we wish to find the set of points, $D_r(c)$, in the complex z plane which satisfy the inequality

$$|P_{r,d_0}(z)| \leq \{cF_r\}^r , \tag{A.1}$$

where $1 \leq c \leq 1/F_r$,

$$P_{r,d_0}(z) = \frac{T_r\left(\dfrac{2z - d_0 - b_0}{d_0 - b_0}\right)}{T_r\left(\dfrac{2 - d_0 - b_0}{d_0 - b_0}\right)} , \tag{A.2}$$

and

$$F_r = \left\{ \frac{1}{T_r\left(\dfrac{2 - d_0 - b_0}{d_0 - b_0}\right)} \right\}^{1/r} . \tag{A.3}$$

From (A.2) and (A.3), the set of points $D_r(c)$ may also be expressed as

$$D_r(c) \equiv \left\{ z : \left| T_r\left(\frac{2z - d_0 - b_0}{d_0 - b_0} \right) \right| \leq c^r \right\} \tag{A.4}$$

We shall consider the special cases of r=1,2 and the limit as r approaches ∞.

<u>CASE 1: r = 1</u>

Since $T_1(s) = s$, we seek those z which satisfy the inequality $\left|\dfrac{2z - d_0 - b_0}{d_0 - b_0}\right| \leq c$. Since $b_0 < d_0, d_0 - b_0 > 0$ and hence we may write

(A.5)
$$\left| z - \left(\frac{d_0 + b_0}{2}\right)\right| \le c\left(\frac{d_0 - b_0}{2}\right).$$

Thus, the set $D_1(c)$ consists of all z on or interior to the circle with center at $\left(0\ ,\ \frac{d_0 + b_0}{2}\right)$ and radius $c\left(\frac{d_0 - b_0}{2}\right)$. Using the maximum modulus theorem, one may easily show that no z exterior to this circle satisfies (A.5).

CASE 2: r = 2

For this case $T_2(s) = 2s^2 - 1$ and thus we seek those z which satisfy

(A.6)
$$\left| 2\left(\frac{2z - d_0 - b_0}{d_0 - b_0}\right)^2 - 1\right| \le c^2 .$$

If we let $a = \frac{d_0 - b_0}{2\sqrt{2}}$, then we may write (A.6) as

$$\left| \frac{1}{a^2}\left(z - \left(\frac{d_0 + b_0}{2}\right)\right)^2 - a^2\right| \le c^2$$

or equivalently

(A.7)
$$\left|\left\{z - \left(\frac{d_0 + b_0}{2}\right) + a\right\}\left\{z - \left(\frac{d_0 + b_0}{2}\right) - a\right\}\right| \le (ac)^2 .$$

If z = x + iy, then (A.7) becomes

(A.8)
$$\sqrt{\left\{\left[x - \left(\frac{d_0 + b_0}{2}\right)\right] + a\right\}^2 + y^2}\sqrt{\left\{\left[x - \left(\frac{d_0 + b_0}{2}\right)\right] - a\right\}^2 + y^2} \le (ac)^2 .$$

The points (x,y) satisfying the equality in (A.8) lie on the ovals of Cassini. See Figure A.1.

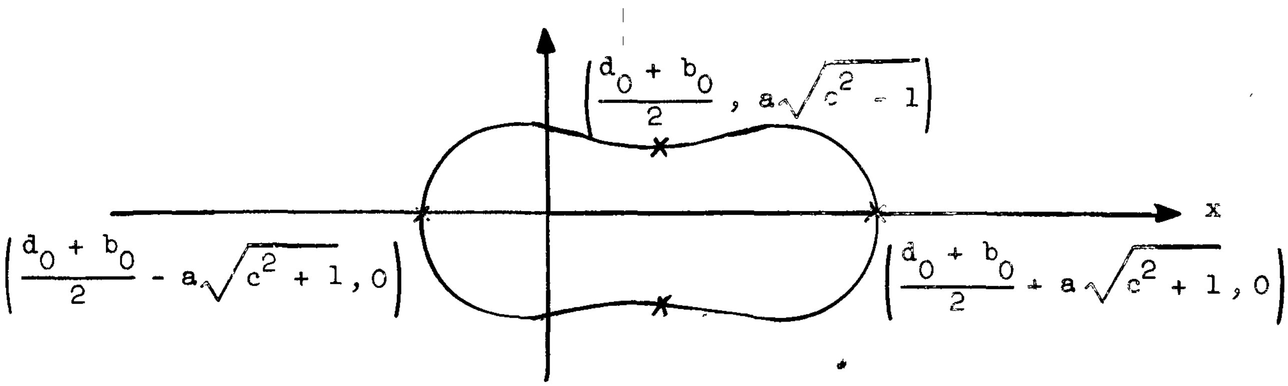


Figure A.1

Thus, again using the maximum modulus theorem, the set $D_2(c)$ consists of all z on or interior to the ovals of Cassini.

CASE 3: $r \to \infty$

If we let

$$\text{(A.9)} \qquad s = \frac{z - \left(\frac{d_0 + b_0}{2}\right)}{\frac{d_0 - b_0}{2}} ,$$

then the Chebyshev polynomial $T_r\left(\frac{2z - d_0 - b_0}{d_0 - b_0}\right)$ may be expressed [Forsythe and Wasow (1960), p. 228] as

$$T_r(s) = \frac{1}{2}\{(s + \sqrt{s^2 - 1})^r + (s - \sqrt{s^2 - 1})^r\}$$

or equivalently as

$$\text{(A.10)} \qquad T_r(s) = \frac{1}{2}\{(s + \sqrt{s^2 - 1})^r + (s + \sqrt{s^2 - 1})^{-r}\} .$$

Now let

$$w = s + \sqrt{s^2 - 1} \ . \quad \text{(A.11)}$$

Thus, $|T_r(s)| \leq c^r$ when $|1/2[w^r + w^{-r}]| \leq c^r$ or equivalently when

$$\left| \left(\frac{w}{c}\right)^r + \frac{1}{(wc)^r} \right| \leq 2 \ . \quad \text{(A.12)}$$

As r approaches ∞ , the inequality (A.12) is satisfied if and only if[1]

$$\left| \frac{w}{c} \right| \leq 1 \ \text{ and } \left| \frac{1}{wc} \right| \leq 1 \ .$$

or equivalently if and only if

$$|w| \leq c \ \text{and} \ |w| \geq = 1/c \ . \quad \text{(A.13)}$$

Hence, w must lie in the closed annulus between the circles $|w| = 1/c$ and $|w| = c$. We now shall proceed to get the corresponding z region.

Solving Eq. (A.11) for s gives

$$s = 1/2\{w + 1/w\} \ . \quad \text{(A.14)}$$

If $c > 1$, then the annulus

---

[1]The if part is obvious. To show the only if part, one need only show that if $\frac{w}{c}$ or $1/|wc|$ is greater than unity, then

$$\lim_{r \to \infty} \left| \left(\frac{w}{c}\right)^r + \left(\frac{1}{wc}\right)^r \right| = \infty \ .$$

$$1/c \leq |w| \leq c$$

in the w plane is mapped [Kober (1957), pg. 62] onto the closed region

$$|s - 1| + |s + 1| \leq \frac{c + 1/c}{2}$$

in the s plane and onto the closed region

$$\text{(A.15)} \qquad |z - d_0| + |z - b_0| \leq \frac{(d_0 - b_0)(c + \frac{1}{c})}{4}$$

in the z plane. See Figure A.2. If $z = x + iy$, then the closed region described by the inequality (A.15) is simply the ellipse

$$\text{(A.16)} \qquad \frac{\left[x - \frac{d_0 + b_0}{2}\right]^2}{\left[\left(\frac{d_0 - b_0}{4}\right)\left(c + \frac{1}{c}\right)\right]^2} + \frac{y^2}{\left[\left(\frac{d_0 - b_0}{4}\right)\left(c - \frac{1}{c}\right)\right]^2} = 1$$

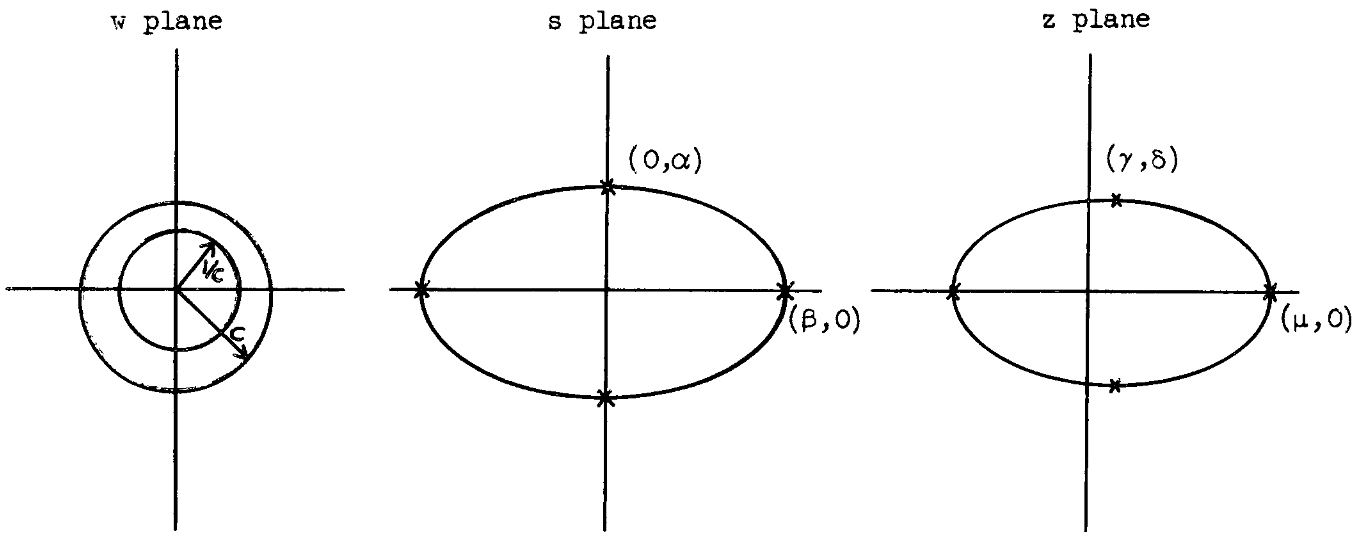
and its interior.

If $c = 1$, the circle $|w| = 1$ is mapped onto the line segment $-1 \leq s \leq 1$ in the s plane and onto the line segment $b_0 \leq z \leq d_0$ in the z plane. Thus, in the limit as r approaches infinity, the region $D_\infty(c)$ consists[1] of all z on or interior to the ellipse (A.16).

The regions $D_1(c)$, $D_2(c)$, and $D_\infty(c)$ for $c = 1$ and $c = 1/F_r$ are given in Figures 2.1, 2.2 and 2.3.

---

[1]A similar type result is given by Wrigley (1963).



$$\alpha = \frac{c - \frac{1}{c}}{2} \qquad \gamma = \frac{d_0 + b_0}{2} \qquad \mu = (d_0 - b_0)\beta + \gamma$$

$$\beta = \frac{c + \frac{1}{c}}{4} \qquad \delta = \frac{d_0 - b_0}{2}\alpha$$

Figure A.2

APPENDIX B

THE INHOMOGENEOUS PROBLEM

In this appendix we shall describe briefly the use of the Chebyshev polynomial method of iteration in the solution of the inhomogeneous matrix problem

$$Ax = s \, , \tag{B.1}$$

where, for convenience, we assume that the nxn nonsingular matrix A is given by A = I - B. We shall also assume that the eigenvalues $\left\{\sigma_i\right\}_{i=1}^{i=n}$ of B are real and are ordered such that

$$|\sigma_n| \le |\sigma_{n-1}| \le \cdots \le |\sigma_2| < \sigma_1 < 1 \tag{B.2}$$

and that the set of eigenvectors for B span the associated vector space.

The matrix problem (B.1) may be solved iteratively using the well known Jacobi method

$$\underline{x}(k+1) = B\underline{x}(k) + \underline{s} \, , \tag{B.3}$$

where k is the iteration index number. If $\underline{E}(k) \equiv \underline{x}(k) - \underline{x}$ is the error vector after k iterations, then it follows from (B.1) and (B.3) that

$$\underline{E}(k) = B\underline{E}(k-1) = B^k\underline{E}(0) \, .$$

Expanding $\underline{E}(0)$ in terms of the eigenvectors of B gives

$$\underline{E}(k) = B^k\left[\sum_i c_i\underline{x}_i\right] = \sum_i (\sigma_i)^k c_i\underline{x}_i \, .$$

Hence, for the Jacobi method of iteration, we see that the most slowly decaying contribution to the initial error vector goes to zero as $(\sigma_1)^k$.

To solve (B.1) using the Chebyshev polynomial method of iteration, one can use [Varga (1962), pg. 138] the iterative procedure

$$\text{(B.4)} \qquad \underline{x}(k+1) = \omega_{k+1}[\underline{v}(k+1) - \underline{x}(k-1)] + \underline{x}(k-1) \; ,$$

where $\underline{v}(k+1) = B\underline{x}(k) + \underline{s}$. The sequence ω_{k+1} is given by $\omega_1 = 1$ and

$$\omega_{k+1} = \frac{2T_k\left(\frac{1}{\sigma_1}\right)}{\sigma_1 T_{k+1}\left(\frac{1}{\sigma_1}\right)} \; , \quad k \geq 1$$

where $T_k(w)$ is the Chebyshev polynomial of degree k in w. The error vector $\underline{E}^{(k)}$ after k Chebyshev iterations can be [Varga (1962)] given by

$$\text{(B.5)} \qquad \underline{E}(k) = P_k(B)\underline{E}(0) \; ,$$

where

$$P_k(w) \equiv \frac{T_k\left(\frac{w}{\sigma_1}\right)}{T_k\left(\frac{1}{\sigma_1}\right)} \; .$$

Thus, the expansion of $\underline{E}(0)$ gives

$$\underline{E}(k) = P_k(B)\left[\sum_i c_i \underline{x}_i\right] = \sum_i P_k(\sigma_i) c_i \underline{x}_i \; .$$

Hence, we see that the most slowly decaying contribution to the initial error vector for the Chebyshev iterations goes to zero as $P_k(\sigma_1)$.

As for the eigenvalue problem, the rate of convergence of the Chebyshev iterations (B.4) is greatly affected by the estimate for σ_1. In practice, two basic approaches are often used to estimate σ_1. One approach [see, for example, Varga (1962), Wachspress (1966), Forsythe and Wasow (1960), and Hageman and Kellogg (1966)] is to obtain an estimate for σ_1 prior to carrying out the Chebyshev iterations. For example, an a priori estimate for σ_1 may be obtained by using the power or Chebyshev iteration method on the matrix B. The second approach is to obtain estimates for σ_1 while carrying out the Chebyshev iterations. In what follows we shall describe a Chebyshev strategy, similar to that given for the eigenvalue problem, which may be used to obtain estimates for σ_1 while carrying out the Chebyshev iterations.

As before, numerical estimates for σ_1 will be obtained by observing the decay rate of the residual vector $\underline{y}(k) \equiv \underline{v}(k) - \underline{x}(k-1)$, where $\underline{v}(k) = B\underline{x}(k-1) + \underline{s}$. We will again use the residual vector quotient

$$Q(k) \equiv \frac{\|\underline{y}(k)\|}{\|\underline{y}(k-1)\|} \tag{B.6}$$

to measure the decay rate and the same Chebyshev strategy as described in Chapter II in obtaining estimates for σ_1.

An initial estimate for σ_1 can be obtained by doing a few Jacobi iterations before starting the Chebyshev iterations. For the Jacobi method we have $\lim_{k\to\infty} Q(k) = \sigma_1$. This follows from the fact that the residual vectors for the Jacobi method satisfy

$$\underline{y}(k) = B\underline{y}(k-1) = B^{k-1}\underline{y}(1) \quad .$$

Suppose now that a Chebyshev polynomial using d_0 as the estimate for σ_1 is started on iteration $k_1 + 1$. Then for iteration $k_1 + r + 1$, $\underline{y}(k_1 + r + 1) = (B - I)\underline{x}(k_1 + r) + \underline{s}$ and since $\underline{s} = (I - B)\underline{x}$ we have

(B.7) $$\underline{y}(k_1 + r + 1) = (B - I)\underline{E}(k_1 + r) \ .$$

Using (B.5) we then get

$$\underline{y}(k_1 + r + 1) = (B - I)P_{r,d_0}(B)\underline{E}(k_1) \ .$$

But from (B.7), $\underline{E}(k_1) = (B - I)^{-1}\underline{y}(k_1 + 1)$ and since $B - I$ commutes with $P_{r,d_0}(B)$ we have

(B.8) $$\underline{y}(k_1 + r + 1) = P_{r,d_0}(B)\underline{y}(k_1 + 1) \ .$$

Thus, for k_1 large enough we have

(B.9) $$Q(k_1 + r + 1) \approx \left| \frac{P_{r,d_0}(\sigma_1)}{P_{r-1,d_0}(\sigma_1)} \right|$$

and

(B.10) $$Q_{r+1} \approx |P_{r,d_0}(\sigma_1)| \ ,$$

where $Q_{r+1} = \prod_{\ell=2}^{r+1} Q(k_1 + \ell)$. Note that Eqs. (B.9) and (B.10) are the same as those obtained for the eigenvalue problem. Thus, as before, one may obtain a new estimate for σ_1 by solving Eq. (B.10) for σ_1.

If the eigenvalues of B are real but now satisfy

$$|\sigma_n| \leq |\sigma_{n-1}| \leq \cdots \leq |\sigma_3| < |\sigma_2| = \sigma_1 < 1 ,$$

where $\sigma_2 = -\sigma_1$, then the above procedure can still be used provided the following changes are noted.

First redefine Q(k) as

$$Q(2k) \equiv \frac{\|\underline{y}(2k)\|}{\|\underline{y}(2(k-1))\|} \quad . \tag{B.6'}$$

Then for the Jacobi iterations we have $\lim_{k\to\infty} Q(2k) = (\sigma_1)^2$ and in place of (B.9) we have

$$Q(k_1 + 2r + 1) \approx \left| \frac{P_{2r,d_0}(\sigma_1)}{P_{2(r-1),d_0}(\sigma_1)} \right| . \tag{B.9'}$$

Equation (B.10) should be replaced by

$$\frac{\|\underline{y}(k_1 + 2r + 1)\|}{\|\underline{y}(k_1 + 1)\|} \approx |P_{2r,d_0}(\sigma_1)| \quad . \tag{B.10'}$$

In the next section we shall consider the special cyclic Chebyshev polynomial method.

1. The Cyclic Chebyshev Method

In this section we assume that the matrix B is written in the special cyclic form $B = \begin{pmatrix} 0 & U \\ L & 0 \end{pmatrix}$ and that Eq. (B.1) is written in the form

$$(B.11) \qquad \begin{pmatrix} I & -U \\ -L & I \end{pmatrix} \begin{pmatrix} \underline{x}_1 \\ \underline{x}_2 \end{pmatrix} = \begin{pmatrix} \underline{s}_1 \\ \underline{s}_2 \end{pmatrix} .$$

The cyclic Chebyshev method of iteration is then defined [Varga (1962), pg. 150] to be

$$(B.12) \quad \begin{cases} \underline{x}_1(k+1) = \alpha_{k+1}[\underline{v}_1(k+1) - \underline{x}_1(k)] + \underline{x}_1(k) \ , \text{ where } \underline{v}_1(k+1) = U\underline{x}_2(k) + \underline{s}_1 \ , \\ \underline{x}_2(k+1) = \beta_{k+1}[\underline{v}_2(k+1) - \underline{x}_2(k)] + \underline{x}_2(k) \ , \text{ where } \underline{v}_2(k+1) = L\underline{x}_1(k+1) + \underline{s}_2 \ . \end{cases}$$

The sequences α_{k+1} and β_{k+1} are given by $\alpha_1 = 1$, $\beta_1 = 2/(2 - \sigma_1^2)$ and for $k \geq 1$

$$\alpha_{k+1} = \frac{1}{1 - \frac{\sigma_1^2}{4}\beta_k} \ , \quad \beta_{k+1} = \frac{1}{1 - \frac{\sigma_1^2}{4}\alpha_{k+1}} \ .$$

It can be shown [Varga (1962)] that the error vectors $\underline{E}_1(k) \equiv \underline{x}_1(k) - \underline{x}_1$ and $\underline{E}_2(k) \equiv \underline{x}_2(k) - \underline{x}_2$ satisfy the equations

$$(B.13) \qquad \underline{E}_1(k) = S_{k-1}(UL)U\underline{E}_2(0) \ , \ \underline{E}_2(k) = R_k(LU)\underline{E}_2(0) \ ,$$

where $S_k(UL)$ and $R_k(LU)$ are polynomials of degree k in UL and LU, respectively, and where $S_0(UL) = R_0(LU) = I$. Also

$$(B.14) \qquad R_k(w^2) = P_{2k}(w) \quad \text{and} \quad wS_k(w^2) = P_{2k+1}(w)$$

where $P_k(w)$ is given by (B.5).

Since $\underline{s}_2 = -L\underline{x}_1 + \underline{x}_2$, the residual vector $\underline{y}_2(k+1) \equiv \underline{v}_2(k+1) - \underline{x}_2(k)$ can be expressed as

$$\underline{y}_2(k+1) = L\underline{x}_1(k+1) - Lx_1 + x_2 - \underline{x}_2(k) = L\underline{E}_1(k+1) - \underline{E}_2(k)$$

and using (B.13) we get

$$\underline{y}_2(k+1) = [LS_k(UL)U - R_k(LU)]\underline{E}_2(0) \tag{B.15}$$

or equivalently since $\underline{E}_2(0) = (LU - I)^{-1}\underline{y}_2(1)$

$$\underline{y}_2(k+1) = [LS_k(UL)U - R_k(LU)][LU - I]^{-1}\underline{y}_2(1) \ .$$

This expression is not quite as nice as the corresponding expression (B.8) for the normal Chebyshev method. However, if d_0 is the estimate for σ_1 used in the generation of the α and β sequences and if $d_0 < \sigma_1$, then for k sufficiently large we still have

$$Q(k+1) \approx \left| \frac{P_{2k,d_0}(\sigma_1)}{P_{2k-2,d_0}(\sigma_1)} \right| ,$$

where $Q(k+1) \equiv ||\underline{y}_2(k+1)||/||\underline{y}_2(k)||$. But now

$$Q_{k+1} \equiv \frac{||\underline{y}_2(k+1)||}{||\underline{y}_2(1)||} \approx f(\sigma_1,d_0)P_{2k,d_0}(\sigma_1) \ ,$$

where $f(\sigma_1,d_0)$ is some function in σ_1 and d_0.

An alternate approach is the following: Let the error vector $\underline{E}_2(k_1)$ at the end of k_1 iterations be given by (B.13), i.e., $\underline{E}_2(k_1) = R_{k_1}(LU)\underline{E}_2(0)$. Now suppose on iteration $k_1 + 1$, a new Chebyshev polynomial is started, i.e., the α and β sequences are started over in (B.12). Then from (B.15) we have

$$\underline{y}_2(k_1 + 1) = (LU - I)\underline{E}_2(k_1) = (LU - I)R_{k_1}(LU)\underline{E}_2(0) \ .$$

But $\underline{E}_2(0) = (LU - I)^{-1}\underline{y}_2(1)$ so that

$$\underline{y}_2(k_1 + 1) = R_{k_1}(LU)\underline{y}_2(1) \ .$$

Hence, under suitable conditions we have

$$\text{(B.16)} \qquad \frac{\|\underline{y}_2(k_1 + 1)\|}{\|\underline{y}_2(1)\|} \approx P_{2k_1,d_0}(\sigma_1) \ .$$

The alternate approach has the disadvantage that the generation of the Chebyshev polynomial must be terminated in order to obtain a new estimate for σ_1. However, the effect of this should be small if proper care is taken.

REFERENCES

1. Bilodeau, G. G. and Hageman, L. A., (1957), "A Survey of Numerical Methods in the Solution of Diffusion Problems," WAPD-TM-64.

2. Clayton, A. J., (1963), "Further Results on Polynomials Having Least Maximum Modulus Over an Ellipse in the Complex Plane," U.K.A.E.A. Memorandum -AEEW- M.348.

3. Faddeev, D. K. and Faddeeva, V. N., (1963), Computational Methods of Linear Algebra, W. H. Freeman and Company, San Francisco and London.

4. Flanders, D. A. and Shortley, G., (1950), "Numerical Determination of Fundamental Modes," J. Appl. Physics, Vol. 21, pages 1326-1332.

5. Forsythe, G. E. and Wasow, W. R., (1960), Finite Difference Methods for Partial Differential Equations, John Wiley and Sons, Inc., New York and London.

6. Hageman, L. A., (1963), "Numerical Methods and Techniques Used in the Two-Dimensional Neutron-Diffusion Program PDQ-5," WAPD-TM-364.

7. Hageman, L. A. and Kellogg, R. B. (1966), "Estimating Optimum Over-relaxation Parameters," WAPD-TM-592.

8. Halmos, P. R., (1957), Introduction to Hilbert Space, Chelsea Publishing Company, New York.

9. Householder, A. S. (1953), Principles of Numerical Analysis, McGraw-Hill Book Co., Inc., New York.

10. Kober, H. (1957), Dictionary of Conformal Representations, Dover Publications, Inc.

11. National Bureau of Standards Applied Mathematics Series .9 (1952), Tables of Chebyshev Polynomials $S_n(x)$ and $c_n(x)$.

12. Perlis, S., (1952), Theory of Matrices, Addison-Wesley Press, Inc. Cambridge.

13. Varga, R. S., (1957), "On Estimating Rates of Convergence in Multigroup Diffusion Problems," WAPD-TM-41.

14. Varga, R. S. (1962), Matrix Iterative Analysis, Prentice-Hall, Inc. Englewood Cliffs, New Jersey.

15. Wachspress, E. L., (1966), Iterative Solution of Elliptic Systems, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

16. Wrigley, H. E., (1963), "Accelerating the Jacobi Method for Solving Simultaneous Equations by Chebyshev Extrapolation when the Eigenvalues of the Iteration Matrix are Complex," Computer Journal, Vol. 6, pages 169-176.

ACKNOWLEDGEMENT

The author would like to thank Dr. R. B. Kellogg for his many comments and suggestions concerning the contents of this paper.

/gb