This is to certify that the

dissertation entitled

Convergence of Several Iterative Methods and Solving Symmetric Tridiagonal Eigenvalue Problems

presented by

Qingchuan Yao

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Applied Mathematics

Major professor

Date June 30, 1998

PLACE IN RETURN BOX
to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE

CONVERGENCE OF SEVERAL ITERATIVE METHODS AND SOLVING SYMMETRIC TRIDIAGONAL EIGENVALUE PROBLEMS

By

Qingchuan Yao

A DISSERTATION

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

DOCTOR OF PHILOSOPHY

Department of Mathematics

1998

ABSTRACT

CONVERGENCE OF SEVERAL ITERATIVE METHODS AND SOLVING SYMMETRIC TRIDIAGONAL EIGENVALUE PROBLEMS

By

Qingchuan Yao

This dissertation studies several iterative methods and presents an algorithm for the eigenvalue problem of the symmetric tridiagonal matrices.

We first propose some modified Halley's iterations for finding the zeros of polynomials. We investigate the non-overshoot properties of the modified Halley iterations and other important properties. We also extend Halley's iteration to systems of polynomial equations in several variables. Then, we study several important properties of the two major concurrent iterative methods: Durand-Kerner's method and Aberth's method for finding all polynomial zeros simultaneously.

Besides Durand-Kerner's method and Aberth's method, several other concurrent iterative methods have been created in the past two decades. However, none of them is of monotonic convergence for solving polynomials with real zeros. The monotonic convergence property plays a key role in solving symmetric tridiagonal eigenvalue problems. Therefore, we propose two new concurrent iterative methods with quadratically convergent rate and cubically convergent rate, respectively. Both of the new methods converge monotonically.

Finally, we present an algorithm for the eigenvalue problem of the symmetric tridiagonal matrices. Our algorithm employs the determinant evaluation, split-and-merge strategy and our newly developed concurrent iterative methods with cubically convergent rate and with monotonic convergence property. Our algorithm is parallel in nature and the preliminary numerical results show that our algorithm is very promising.

To: Jenny Yao ——— my lovely daughter

Lianfen Qian —— my beloved wife

ACKNOWLEDGMENTS

I would like to thank Professor Tien-Yien Li, my dissertation advisor, for his constant encouragement and support during my graduate study at Michigan State University. I would also like to thank him for suggesting the problem and the helpful directions which made this work possible.

I would like to thank Professor Qiang Du from whom I took several courses which laid the foundation for this research. I would also like to thank him for the helpful discussion.

I would like to thank my dissertation committee members Professor Wei-Eihn Kuan, Professor Jay C. Kurtz, Professor William T. Sledd, and Professor Zhengfang Zhou for their valuable suggestions and their time.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In this work, we deal with finding zeros of polynomials and solving matrix eigenvalue problems.

To approximate a zero of a function $f(z)$, real or complex, Halley's iteration is defined as follows:

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k) - \frac{1}{2}\frac{f(z_k)f''(z_k)}{f'(z_k)}}, \quad k = 0, 1, 2, \ldots,$$

which was initially discovered in 1694 by E. Halley, who computed the orbit of the Halley comet. Halley's iteration is cubically convergent for simple zeros of $f(z)$. However, it converges only linearly for multiple zeros.

In Chapter 2, we apply Newton's iteration to derive the following modified Halley iteration:

$$z_{k+1} = z_k - \frac{f(z_k)}{\frac{1+m}{2m}f'(z_k) - \frac{1}{2}\frac{f(z_k)f''(z_k)}{f'(z_k)}}, \quad k = 0, 1, 2, \ldots,$$

which approximates the multiple zeros with multiplicity m of $f(z)$ with cubic convergence rate. We show the *non-overshoot* properties of this modified iteration for polynomial with real zeros and obtain the monotonic convergence of the Halley iteration in this situation as a by-product. Moreover, this modified Halley iteration will be used to derive an inequality which gives a circle containing at least one zero of the polynomial. This inequality is at least as good as Kahan's inequality derived by using the modified Laguerre iteration. We will also study the Halley iteration in Banach space and establish a convergence theorem with an optimal error bound similar in spirit to the Newton-Kantorovich theorem for Newton's iteration.

In the past two decades the so-called concurrent iterative methods for complex polynomial $P(z)$ of degree n have been extensively studied. Two major concurrent iterative methods among them are the Durand-Kerner method:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{P(z_i^{(k)})}{\prod_{j=1,j\neq i}^{n}(z_i^{(k)} - z_j^{(k)})}, \quad k = 0,1,2,\ldots,$$

and the Aberth method:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{1}{\frac{P'(z_i^{(k)})}{P(z_i^{(k)})} - \sum_{j=1,j\neq i}^{n} \frac{1}{z_i^{(k)} - z_j^{(k)}}}, \quad k = 0,1,2,\ldots,$$

for $i = 1,2,\ldots,n$. The attractive feature of Durand-Kerner's method is its quadratic convergence with no derivative evaluations and Aberth's method is remarkable for its cubic convergence with no second derivative evaluations. Furthermore, they both converge to all the n zeros of $P(z)$ simultaneously when the polynomial has only simple zeros.

In Chapter 3, we first present a new derivation of Durand-Kerner's method by homotopy. The new derivation gives a geometric interpretation of this method. We then propose the following two-step iterative scheme:

$$z_i^{(k+2)} = z_i^{(k+1)} - (z_i^{(k+1)} - z_i^{(k)}) \left[\sum_{j=1,j\neq i}^{n} \frac{z_j^{(k)} - z_j^{(k+1)}}{z_i^{(k+1)} - z_j^{(k)}} \right] \left[\prod_{j=1,j\neq i}^{n} \frac{z_i^{(k+1)} - z_j^{(k)}}{z_i^{(k+1)} - z_j^{(k+1)}} \right],$$

$k = 0,1,2,\ldots,$ for $i = 1,2,\ldots,n$ with given $z_1^{(0)}, z_2^{(0)},\ldots,z_n^{(0)}$ and

$$z_i^{(1)} = z_i^{(0)} - \frac{P(z_i^{(0)})}{\prod_{j=1,j\neq i}^{n}(z_i^{(0)} - z_j^{(0)})}, \quad i = 1,2,\ldots,n.$$

We will show that this two-step iterative scheme is equivalent to Durand-Kerner's method. Notice that our new scheme requires no polynomial evaluations. Similarly, we will show that Aberth's method is equivalent to:

$$z_i^{(k+1)} = z_i^{(k)} - \frac{d_i^{(k)}}{1 - \sum_{j=1,j\neq i}^{n} \frac{d_j^{(k)}}{z_j^{(k)} - z_i^{(k)}}}, \quad k = 0,1,2,\ldots; \quad i = 1,2,\ldots,n$$

where

$$d_i^{(k)} = \frac{P(z_i^{(k)})}{\prod_{j=1,j\neq i}^{n}(z_i^{(k)} - z_j^{(k)})}, \quad k = 0,1,2,\ldots; \quad i = 1,2,\ldots,n$$

with no derivative evaluations.

In Chapter 4, we propose some new concurrent iterative methods with quadratic convergence rate without derivative evaluations. Those methods converge monotonically when they are used to approximate real zeros of a complex polynomial $P(z)$ with degree n. The main idea is to define a pair of sequences $\{x_i^{(k)}\}_{i=1}^{n}$ and $\{y_i^{(k)}\}_{i=1}^{n}$ by means of the following iterations:

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k)})\prod_{j=i+1}^{n}(x_i^{(k)} - y_j^{(k)})}, \quad k = 0,1,2,\ldots, \\ y_i^{(k+1)} &= y_i^{(k)} - \frac{P(y_i^{(k)})}{\prod_{j=1}^{i-1}(y_i^{(k)} - x_j^{(k)})\prod_{j=i+1}^{n}(y_i^{(k)} - y_j^{(k)})}, \quad k = 0,1,2,\ldots, \end{aligned}$$

and $i = 1,2,\ldots,n$. We first establish the results of these concurrent iterations for the simple zeros of $P(z)$. For the real multiple zeros as well as a cluster of zeros of $P(z)$, we propose the corresponding new modified concurrent iterative methods which converge monotonically. We also prove the non-overshoot properties and the quadratic convergence rate of the iterations.

One of our main purposes in studying the new concurrent iterative methods with monotonic convergence is to solve the symmetric eigenvalue problems in parallel. Finding the eigenvalues of an $n \times n$ symmetric tridiagonal matrix A is equivalent to solving $\det[A - \lambda I] = 0$, a polynomial of a single variable λ with real zeros and with degree n. In Chapter 5, we present an algorithm for the eigenvalue problem of symmetric tridiagonal matrices by using the following new concurrent iterations

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} - \frac{n_i P(x_i^{(k)})}{P'(x_i^{(k)}) - P(x_i^{(k)})\sum_{j=1,j\neq i}^{m}\frac{n_j}{x_i^{(k)} - y_j^{(k)}}}, \quad k = 0,1,2,\ldots, \\ y_i^{(k+1)} &= y_i^{(k)} - \frac{n_i P(y_i^{(k)})}{P'(y_i^{(k)}) - P(y_i^{(k)})\sum_{j=1,j\neq i}^{m}\frac{n_j}{y_i^{(k)} - x_j^{(k)}}}, \quad k = 0,1,2,\ldots, \end{aligned}$$

where $P(z) = \prod_{j=1}^{m}(z - \lambda_j)^{n_j}$, $\sum_{j=1}^{m} n_j = n$ and $i = 1,2,\ldots,m$. We show that the above concurrent iterations are of monotonic convergence for real multiple zeros

of $P(z)$ with non-overshoot properties for clusters of real zeros of $P(z)$. Moreover, these iterations are cubically convergent. The algorithm is applied to extract the eigenvalues of symmetric tridiagonal matrices iteratively. With promising numerical results and its natural parallelism our algorithm is therefore an excellent candidate for advanced architectures.

Chapter 2

The Halley Method

2.1 Introduction

The monotonically convergent Laguerre iteration and modified Laguerre iteration with cubic convergence rate for solving equations of polynomials with real zeros have a very important application in solving symmetric tridiagonal eigenvalue problems, which have only real eigenvalues. Comprehensive numerical results in this direction can be found in Li & Zeng [36] and K. Li, T.Y. Li & Z. Zeng [35]. In Laguerre's iteration or modified Laguerre's iteration, one needs to take a square root for each iteration step. In actual computation, because of computer round-off errors, the values inside the square root may become negative. To avoid this problem, in this chapter we discuss iterations called Halley's iteration and modified Halley's iteration without any square roots in their iteration formulae. We will show that Halley's iteration and modified Halley's iteration share the same properties as Laguerre's iteration and modified Laguerre's iteration when they are used to find the zeros of polynomials with real zeros. Like Newton's iteration, Halley's iteration has a very natural extension to systems of polynomial equations in several variables or even to nonlinear operators in real or complex Banach space. To the best of our knowledge, it is still not clear that the Laguerre iteration can be carried out in the same manner [30].

In Section 2.2 we apply Newton's iteration to derive the modified Halley iteration, which is suitable for finding approximations to multiple roots of $f(x) = 0$ with cubic convergence rate. In Section 2.3 we exhibit the non-overshoot properties of this iteration and obtain the monotonic convergence of Halley's iteration for polynomials with real zeros as a by-product. In Section 2.4 a circle is given by modified Halley's iteration that contains at least one zero of the polynomial. In Section 2.5 this iteration is converted to another iteration called two-sided modified Halley's iteration. The non-overshoot properties of this iteration when it is used for polynomials with real zeros is investigated.

In Section 2.6 we establish a convergence theorem for Halley's iteration in Banach space with an optimal error bound similar to the Newton-Kantorovich Theorem for Newton's iteration (see Gragg and Tapia [24]). From the optimal error bound we obtained, it can be easily seen that the convergence rate of Halley's iteration in Banach space is cubic.

2.2 Derivation for the modified Halley iteration and its convergence rate

Consider the problem of approximating a solution of a real or complex equation

$$f(x) = 0. \tag{2.1}$$

The formula

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k) - \frac{1}{2}\frac{f(x_k)f''(x_k)}{f'(x_k)}} \equiv H(x_k)\ , \quad k = 0, 1, 2, \ldots, \tag{2.2}$$

is known as Halley's iteration, see Bateman [5], Frame [19, 20], Stewart [50], Kiss [33], Snyder [48], Šafiev [44], Traub [51], Davies & Dawson [10], Brown [8], Hansen & Patrick [27], Popovski [42], Alefeld [3], Zheng [58], Gander [21], Scavo & Thoo [46]. E. Halley, who computed the orbit of Halley's comet, discovered a special case of this formula in 1694. Scavo and Thoo [46] have shown that there are many ways to derive Halley's iteration (2.2).

Now rewrite $H(x)$ as

$$H(x) = x - \frac{2f(x)f'(x)}{2(f'(x))^2 - f(x)f''(x)} , \tag{2.3}$$

then

$$H'(x) = -\frac{1}{2}(H(x) - x)^2 \left(\frac{f'''(x)}{f'(x)} - \frac{3}{2} \left(\frac{f''(x)}{f'(x)} \right)^2 \right) . \tag{2.4}$$

If z is a simple zero of $f(x)$, we can easily see, from (2.4) above, that $H'(z) = H''(z) = 0$ and $H'''(z) \neq 0$. Thus, Halley's iteration (2.2) converges cubically in this case.

However, if z is a zero of $f(x)$ with multiplicity $m > 1$, then $f(x) = (x - z)^m h(x)$ and $h(z) \neq 0$. Thus,

$$\begin{aligned} f'(x) &= (x - z)^{m-1}\psi(x), \\ f''(x) &= (x - z)^{m-2}[(m - 1)\psi(x) + (x - z)\psi'(x)] \end{aligned}$$

where $\psi(x) = mh(x) + (x - z)h'(x)$. Substituting these into (2.3) yields

$$H(x) = x - (x - z)\Psi(x)$$

where

$$\Psi(x) = \frac{2h(x)\psi(x)}{2\psi(x)^2 - h(x)[(m - 1)\psi(x) + (x - z)\psi'(x)]} .$$

Now $H'(x) = 1 - (x - z)\Psi'(x) - \Psi(x)$, so

$$H'(z) = 1 - \Psi(z) = 1 - \frac{2}{m + 1} = \frac{m - 1}{m + 1} \tag{2.5}$$

which implies $0 < H'(z) < 1$ for all $m > 1$. Thus, Halley's iteration (2.2) converges only linearly for a multiple root of $f(x) = 0$.

Suppose $f(x) = (x - z)^m h(x)$ where z is a zero of $f(x)$ with multiplicity m and $h(z) \neq 0$. Let

$$g_m(x) = \frac{f(x)}{[f'(x)]^{\frac{m}{m+1}}} .$$

Then

$$g_m(x) = (x - z)^{\frac{2m}{1+m}} \cdot t(x) \tag{2.6}$$

where

$$t(x) = \frac{h(x)}{[mh(x) + (x - z)h'(x)]^{\frac{m}{1+m}}} . \tag{2.7}$$

Now applying Newton's iteration with multiplicity $\frac{2m}{1+m}$ to the equation $g_m(x) = 0$, we have

$$x_{k+1} = x_k - \frac{2m}{1+m} \cdot \frac{g_m(x_k)}{g'_m(x_k)}, \quad k = 0.1, 2, \ldots . \tag{2.8}$$

Since

$$g'_m(x) = \frac{f'(x) - \frac{m}{1+m} \cdot \frac{f(x)f''(x)}{f'(x)}}{[f'(x)]^{\frac{m}{1+m}}},$$

we obtain the following **Modified Halley Iteration**

$$x_{k+1} = x_k - \frac{f(x_k)}{\frac{1+m}{2m} f'(x_k) - \frac{1}{2} \frac{f(x_k)f''(x_k)}{f'(x_k)}} \equiv H_m(x_k) \tag{2.9}$$

where

$$H_m(x) = x - \frac{f(x)}{\frac{1+m}{2m} f'(x) - \frac{1}{2} \frac{f(x)f''(x)}{f'(x)}} . \tag{2.10}$$

When $m = 1$, this modified Halley's iteration (2.9) is exactly the Halley iteration in (2.2). Hansen and Patrick [27] have derived a one-parameter family of iterations for finding the multiple zeros of $f(x)$. Their one-parameter family includes the modified Halley iteration (2.9) as a special case, but with our derivation, it is easier to show the cubic convergence of this iteration when it is used to approximate an m–fold root of $f(x) = 0$.

Theorem 2.2.1 *Modified Halley's iteration (2.9) converges cubically to an m-fold root of* $f(x) = 0$.

Proof For $f(x) = (x - z)^m h(x)$ with $h(z) \neq 0$, we have

$$H_m(x) = x - \frac{2m}{1+m} \cdot \frac{g_m(x)}{g'_m(x)}.$$

Now,

$$H'_m(x) = 1 - \frac{2m}{1+m} + \frac{2m}{1+m} \cdot \frac{g_m(x)g''_m(x)}{[g'_m(x)]^2}$$

and

$$H''_m(x) = \frac{2m}{1+m} \cdot \frac{[g'_m(x)g''_m(x) + g_m(x)g'''_m(x)][g'_m(x)]^2 - 2g_m(x)g'_m(x)[g''_m(x)]^2}{[g'_m(x)]^4}$$

where

$$
\begin{aligned}
g'_m(x) &= \left(\frac{2m}{1+m}\right)(x-z)^{\frac{m-1}{m+1}} \cdot t(x) + (x-z)^{\frac{2m}{1+m}} \cdot t'(x) \; , \\
g''_m(x) &= \left(\frac{2m}{1+m}\right)\left(\frac{m-1}{m+1}\right)(x-z)^{\frac{-2}{m+1}} \cdot t(x) \\
&\quad + \left(\frac{4m}{1+m}\right)(x-z)^{\frac{m-1}{m+1}} \cdot t'(x) + (x-z)^{\frac{2m}{1+m}} \cdot t''(x)
\end{aligned}
$$

and

$$
\begin{aligned}
g'''_m(x) &= \left(\frac{2m}{1+m}\right)\left(\frac{m-1}{m+1}\right)\left(\frac{-2}{m+1}\right)(x-z)^{\frac{-3-m}{m+1}} \cdot t(x) \\
&\quad + \left(\frac{2m}{1+m}\right)\left(\frac{m-1}{m+1}\right)(x-z)^{\frac{-2}{m+1}} \cdot t'(x) \\
&\quad + \left(\frac{4m}{1+m}\right)\left(\frac{m-1}{m+1}\right)(x-z)^{\frac{-2}{m+1}} \cdot t'(x) + O((x-z)^{\frac{m-1}{m+1}}) \; .
\end{aligned}
$$

It follows that

$$\lim_{x \to z} H'_m(x) = 0$$

and

$$\lim_{x \to z} H''_m(x) = 0$$

because $t(z) \neq 0$ and $t'(z) = 0$. □

Remark 2.2.1 *Let*

$$H_c(x) = x - \frac{2f(x)f'(x)}{c(f'(x))^2 - f(x)f''(x)}, \tag{2.11}$$

then

$$H'_c(x) = -\frac{1}{2}(H_c(x) - x)^2 \left(\frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2 - \frac{c^2 - 2c}{2}\left(\frac{f'(x)}{f(x)}\right)^2 \right).$$

It follows that $H'_c(z) = H''_c(z) = 0$ and $H'''_c(z) \neq 0$ only if $c = \frac{m+1}{m}$ and z is a zero of $f(x)$ with multiplicity m. This seems to be a better reason for the cubic convergence of the modified Halley iteration.

If the actual multiplicity M of the zero z of $f(x)$ is known, then the presumed multiplicity m appeared in (2.9) can be taken as M. Theorem 2.2.1 above asserts that the convergence of modified Halley's iteration (2.9) is cubic in this situation. In actual computations, it is often the case that the actual multiplicity M of the zero

z of $f(x)$ is not known a priori, so the presumed multiplicity m in (2.9) is usually different from the actual multiplicity M. In this case, the following theorem suggests that the convergence of the modified Halley iteration (2.9) is only linear.

Theorem 2.2.2 *If $m \neq M$, then*

$$H'_m(z) = \frac{M-m}{M+m}$$

where $H_m(x)$ is the modified Halley iteration function (2.10) and z is the zero of $f(x)$ with multiplicity M.

Proof Let $f(x) = (x-z)^M h(x)$ where $h(z) \neq 0$, then

$$H_m(x) = x - (x-z)\Psi(x)$$

where

$$\Psi(x) = \frac{2h(x)\psi(x)}{\frac{1+m}{m}\psi(x)^2 - h(x)[(M-1)\psi(x) + (x-z)\psi'(x)]}$$

with $\psi(x) = Mh(x) + (x-z)h'(x)$. Thus,

$$H'_m(z) = 1 - \Psi(z) = 1 - \frac{2m}{M+m} = \frac{M-m}{M+m}.$$

□

2.3 Non-overshoot properties and monotonic convergence

In this section we will show that the modified Halley iteration (2.9) can not skip over m zeros or a zero with multiplicity greater than or equal to m for polynomials with only real zeros. With this property the monotonic convergence of Halley's iteration (2.2) is obtained as a by-product.

Theorem 2.3.1 *Let f be a polynomial with real zeros $\lambda_1 < \lambda_2 < \cdots < \lambda_n$. Then, for any $x \in (\lambda_i, \lambda_{i+1})$ and any positive integer m, we have*

$$\lambda_i < x < H_1(x) < H_2(x) < \cdots < H_m(x) < \lambda_{i+m}\ , \quad \text{if } \frac{f(x)}{f'(x)} < 0 \qquad (2.12)$$

and

$$\lambda_{i-m+1} < H_m(x) < H_{m-1}(x) < \cdots < H_1(x) < x < \lambda_{i+1} \ , \quad \textit{if} \ \frac{f(x)}{f'(x)} > 0 \qquad (2.13)$$

with the convention $\lambda_i = -\infty$ *for* $i \leq 0$ *and* $\lambda_i = +\infty$ *for* $i \geq n+1$.

Proof Since

$$\frac{f'(x)}{f(x)} = \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)} \ ,$$

we have

$$-\left(\frac{f'(x)}{f(x)}\right)' = \frac{(f')^2 - f \cdot f''}{f^2} = \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)^2} \ .$$

Rewrite the modified Halley iteration function as

$$\begin{aligned} H_m(x) &= x - \frac{2mf(x)f'(x)}{(m+1)[f'(x)]^2 - mf(x)f''(x)} \\ &= x - \frac{1}{F_m(x)} \end{aligned}$$

where

$$F_m(x) = \frac{(m+1)[f'(x)]^2 - mf(x)f''(x)}{2mf(x)f'(x)},$$

or,

$$\begin{aligned} F_m(x) &= \frac{[f'(x)]^2 + m[(f'(x))^2 - f(x)f''(x)]}{2m(f(x))^2} \cdot \frac{f(x)}{f'(x)} \\ &= \frac{1}{2} \cdot \left[\frac{1}{m} \cdot \frac{f'}{f} + \frac{f}{f'} \cdot \frac{(f')^2 - ff''}{(f)^2}\right] \\ &= \frac{1}{2} \cdot \left[\frac{1}{m} \cdot \frac{f'}{f} + \frac{f}{f'} \cdot \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)^2}\right] \ . \end{aligned}$$

We have the following two cases:

1. If $\frac{f(x)}{f'(x)} < 0$, then

$$\begin{aligned} -F_m(x) &= \frac{1}{2} \cdot \left[\frac{1}{m} \cdot \frac{-f'}{f} + \frac{-f}{f'} \cdot \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)^2}\right] \\ &> \frac{1}{2} \cdot \left[\frac{1}{m} \cdot \frac{-f'}{f} + \frac{-f}{f'} \cdot \sum_{j=i+1}^{i+m} \frac{1}{(x-\lambda_j)^2}\right] \end{aligned}$$

$$\begin{aligned}
&> \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{-f'}{f}+\frac{-f}{f'}\cdot\frac{m}{(x-\lambda_{i+m})^2}\right]\\
&\geq \frac{1}{|x-\lambda_{i+m}|} \qquad (\text{ since } \frac{1}{2}(a+b)\geq\sqrt{ab}\ )\\
&= \frac{1}{\lambda_{i+m}-x}\\
&> 0.
\end{aligned}$$

So

$$x < H_m(x) = x - \frac{1}{F_m(x)} < x + \lambda_{i+m} - x = \lambda_{i+m}\ .$$

2. If $\frac{f(x)}{f'(x)} > 0$, then

$$\begin{aligned}
F_m(x) &= \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'}{f}+\frac{f}{f'}\cdot\sum_{j=1}^{n}\frac{1}{(x-\lambda_j)^2}\right]\\
&> \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'}{f}+\frac{f}{f'}\cdot\sum_{j=i-m+1}^{i}\frac{1}{(x-\lambda_j)^2}\right]\\
&> \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'}{f}+\frac{f}{f'}\cdot\frac{m}{(x-\lambda_{i-m+1})^2}\right]\\
&\geq \frac{1}{|x-\lambda_{i-m+1}|} \qquad (\text{ since } \frac{1}{2}(a+b)\geq\sqrt{ab}\ )\\
&= \frac{1}{x-\lambda_{i-m+1}}\\
&> 0.
\end{aligned}$$

So

$$x > H_m(x) = x - \frac{1}{F_m(x)} > x - (x-\lambda_{i-m+1}) = \lambda_{i-m+1}\ .$$

By 1 and 2 above, and

$$\begin{aligned}
\frac{\partial H_m(x)}{\partial m} &= \frac{-2f(x)(f'(x))^3}{m^2\cdot\left[\frac{m+1}{m}\cdot(f'(x))^2-f(x)f''(x)\right]^2}\\
&= \begin{cases} >0 & \text{if } \frac{f(x)}{f'(x)}<0,\\ <0 & \text{if } \frac{f(x)}{f'(x)}>0,\end{cases}
\end{aligned}$$

the results in (2.12) and (2.13) follow. □

Remark 2.3.1 *From Theorem 2.3.1, the modified Halley iteration (2.9) can not overshoot as many as m zeros of polynomial f.*

Remark 2.3.2 *When $f(x)$ is a polynomial with zeros*

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n,$$

then the zeros of $f'(x)$

$$\lambda'_1 < \lambda'_2 < \cdots < \lambda'_{n-1}$$

have the following property:

$$\lambda_1 < \lambda'_1 < \lambda_2 < \lambda'_2 < \cdots < \lambda_i < \lambda'_i < \lambda_{i+1} < \cdots < \lambda'_{n-1} < \lambda_n.$$

Since

$$-\left(\frac{f'(x)}{f(x)}\right)' = \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)^2} > 0 ,$$

$\frac{f'(x)}{f(x)}$ is monotonically decreasing in any interval $(\lambda_i, \lambda_{i+1})$. So $\frac{f(x)}{f'(x)}$ is monotonically increasing in intervals (λ_i, λ'_i) and $(\lambda'_i, \lambda_{i+1})$. Thus, $\frac{f(x)}{f'(x)}$ is negative for all x in $(\lambda'_i, \lambda_{i+1}) \cup (-\infty, \lambda_1)$ and positive for all x in $(\lambda_i, \lambda'_i) \cup (\lambda_n, +\infty)$. Therefore, by taking $m = 1$ in Theorem 2.3.1, we have the following monotonic convergence of Halley's iteration (2.2).

Corollary 2.3.1 *Let $x_{k+1} = H_1(x_k)$, $k = 0, 1, 2, \ldots$, then for the polynomial $f(x)$ as in Theorem 2.2.2, the following two cases hold:*

Case 1 *For any $x_0 \in (\lambda'_i, \lambda_{i+1})$, we have*

$$x_0 < x_1 < x_2 < \cdots < x_k < \cdots < \lambda_{i+1} \tag{2.14}$$

and $\{x_k\}_{k=0}^{\infty}$ converges to λ_{i+1} cubically.

Case 2 *For any $x_0 \in (\lambda_i, \lambda'_i)$, we have*

$$\lambda_i < \cdots < x_k < \cdots < x_2 < x_1 < x_0 \tag{2.15}$$

and $\{x_k\}_{k=0}^{\infty}$ converges to λ_i cubically.

For Halley's iteration (2.2), Davies and Dawson [10] gave a proof of the monotonic convergence when the iteration is applied to approximate the zeros of the entire function

$$h(x) = x^m \exp(a + bx - cx^2) \prod_{i=1}^{\infty} \left(1 - \frac{x}{\alpha_i}\right) e^{\frac{x}{\alpha_i}}$$

where m is a non-negative integer, a, b, c are real numbers with $c \geq 0$ and α_i are real numbers where $\sum_{i=1}^{\infty} \alpha_i^{-2}$ is convergent.

Theorem 2.3.2 *Suppose that*

$$f(x) = c(x-\lambda)^m \cdot \prod_{j=1}^{n-m} (x - \lambda_j)$$

where c is a constant and

$$\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_i < \lambda < \lambda_{i+1} \leq \cdots \leq \lambda_{n-m} \ ,$$

and suppose $f'(x)$ has zeros λ'_i and λ'_{i+1} with

$$\lambda_i < \lambda'_i < \lambda < \lambda'_{i+1} < \lambda_{i+1}.$$

Let

$$x_{k+1} = H_m(x_k), \ k = 0, 1, 2, \ldots,$$

then either one of the following two cases holds:

Case 1 *For any $x_0 \in (\lambda'_i, \lambda)$,*

$$\lambda'_i < x_0 < x_1 < x_2 < \cdots < x_k < \cdots < \lambda \tag{2.16}$$

and $\{x_k\}_{k=0}^{\infty}$ converges to λ cubically.

Case 2 *For any $x_0 \in (\lambda, \lambda'_{i+1})$,*

$$\lambda < \cdots < x_k < \cdots < x_2 < x_1 < x_0 < \lambda'_{i+1} \tag{2.17}$$

and $\{x_k\}_{k=0}^{\infty}$ converges to λ cubically.

Proof Since

$$\frac{f'(x)}{f(x)} = \frac{m}{x-\lambda} + \sum_{j=1}^{n-m} \frac{1}{x-\lambda_j},$$

we have

$$-\left(\frac{f'(x)}{f(x)}\right)' = \frac{(f')^2 - ff''}{f^2} = \frac{m}{(x-\lambda)^2} + \sum_{j=1}^{n-m} \frac{1}{(x-\lambda_j)^2} > 0.$$

So $\frac{f(x)}{f'(x)}$ is negative for all x in (λ'_i, λ) and positive for all x in $(\lambda, \lambda'_{i+1})$. On the other hand,

$$H_m(x) = x - \frac{1}{F_m(x)}$$

where

$$\begin{aligned}
F_m(x) &\equiv \frac{(m+1)[f'(x)]^2 - mf(x)f''(x)}{2mf(x)f'(x)} \\
&= \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'}{f} + \frac{f}{f'}\cdot\frac{(f')^2 - ff''}{f^2}\right] \\
&= \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'}{f} + \frac{f}{f'}\left(\frac{m}{(x-\lambda)^2} + \sum_{j=1}^{n-m}\frac{1}{(x-\lambda_j)^2}\right)\right].
\end{aligned}$$

Hence, we have the following two cases:

1. If $x_0 \in (\lambda_i', \lambda)$, then $\frac{f(x_0)}{f'(x_0)} < 0$. So

$$\begin{aligned}
-F_m(x_0) &= \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{-f'(x_0)}{f(x_0)} + \frac{-f(x_0)}{f'(x_0)}\cdot\left(\frac{m}{(x_0-\lambda)^2} + \sum_{j=1}^{n-m}\frac{1}{(x_0-\lambda_j)^2}\right)\right] \\
&> \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{-f'(x_0)}{f(x_0)} + \frac{-f(x_0)}{f'(x_0)}\cdot\frac{m}{(x_0-\lambda)^2}\right] \\
&> \frac{1}{|x_0-\lambda|} \\
&= \frac{1}{\lambda - x_0} \\
&> 0.
\end{aligned}$$

Thus,

$$\lambda_i' < x_0 < x_1 = H_m(x_0) = x_0 - \frac{1}{F_m(x_0)} < \lambda.$$

2. If $x_0 \in (\lambda, \lambda_{i+1}')$, then $\frac{f(x_0)}{f'(x_0)} > 0$. So

$$\begin{aligned}
F_m(x_0) &= \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'(x_0)}{f(x_0)} + \frac{f(x_0)}{f'(x_0)}\cdot\left(\frac{m}{(x_0-\lambda)^2} + \sum_{j=1}^{n-m}\frac{1}{(x_0-\lambda_j)^2}\right)\right] \\
&> \frac{1}{2}\cdot\left[\frac{1}{m}\cdot\frac{f'(x_0)}{f(x_0)} + \frac{f(x_0)}{f'(x_0)}\cdot\frac{m}{(x_0-\lambda)^2}\right] \\
&> \frac{1}{|x_0-\lambda|} \\
&= \frac{1}{x_0 - \lambda} \\
&> 0.
\end{aligned}$$

Thus,

$$x_0 > x_1 = H_m(x_0) = x_0 - \frac{1}{F_m(x_0)} > \lambda.$$

□

Remark 2.3.3 *By Theorem 2.3.2, the modified Halley iteration can not overshoot a zero λ with multiplicity greater than or equal to m.*

2.4 Over-estimation of the errors by the modified Halley iteration

For a polynomial f of degree n, the following two inequalities can be found in [29, 30]:
Laguerre's inequality: A zero z of f closest to an approximation x satisfies

$$|z - x| \leq n|N(x) - x| = n\left|\frac{f(x)}{f'(x)}\right| \equiv NB \tag{2.18}$$

where $N(x) = x - \frac{f(x)}{f'(x)}$ is the Newton iteration function.
Kahan's inequality: A zero z of f closest to an approximation x satisfies

$$|z - x| \leq \sqrt{\frac{n}{m}} \cdot |L_m^{\pm}(x) - x| \equiv LB \tag{2.19}$$

where

$$L_m^{\pm}(x) = x + \frac{n}{\left(-\frac{f'(x)}{f(x)}\right) \pm \sqrt{\frac{n-m}{m}\left[(n-1)\left(-\frac{f'(x)}{f(x)}\right)^2 - n\left(\frac{f''(x)}{f(x)}\right)\right]}} \tag{2.20}$$

which is the modified Laguerre iteration function for multiple zeros of $f(x)$.

These inequalities estimate the errors of approximations to a zero z of a polynomial f of degree n. In this spirit, the following theorem provides a similar inequality for the modified Halley iteration.

Theorem 2.4.1 *A zero z of a polynomial $f(x)$ of degree n closest to an approximation x satisfies*

$$|z - x| \leq \sqrt{\frac{n^2 + nm}{2m} \cdot \left|\frac{f(x)}{f'(x)}\right| \cdot |H_m(x) - x|} \equiv HB \tag{2.21}$$

where H_m is the modified Halley iteration function in (2.10).

Remark 2.4.1 *The above theorem ensures that for the modified Halley iteration*

$$x_{k+1} = H_m(x_k), \quad k = 0, 1, 2, \ldots,$$

at least one zero of the polynomial $f(x)$ is contained in each circle

$$|z - x_k| \leq \sqrt{\frac{n^2 + nm}{2m} \cdot \left|\frac{f(x_k)}{f'(x_k)}\right| \cdot |x_{k+1} - x_k|}\ , \quad k = 0, 1, 2, \ldots.$$

Proof of Theorem 2.4.1 Let

$$f(x) = c(x - z_1)(x - z_2)\cdots(x - z_n)$$

where c is a complex constant. Then

$$\frac{f'(x)}{f(x)} = \sum_{j-1}^{n} \frac{1}{x - z_j}$$

and

$$\frac{(f')^2 - ff''}{f^2} = \sum_{j=1}^{n} \frac{1}{(x - z_j)^2}\ .$$

Since

$$H_m(x) = x - \frac{1}{F_m(x)}$$

where

$$F_m(x) = \frac{(m+1)[f'(x)]^2 - mf(x)f''(x)}{2mf(x)f'(x)}\ ,$$

$$|H_m(x) - x| = \frac{1}{|F_m(x)|}\ .$$

Now,

$$\begin{aligned}
|F_m(x)| &= \left|\frac{1}{2} \cdot \left[\frac{1}{m} \cdot \left(\frac{f'}{f}\right)^2 + \frac{(f')^2 - ff''}{f^2}\right]\right| \cdot \left|\frac{f(x)}{f'(x)}\right| \\
&= \left|\frac{1}{2} \cdot \left[\frac{1}{m} \cdot \left(\sum_{j=1}^{n} \frac{1}{x - z_j}\right)^2 + \sum_{j=1}^{n} \frac{1}{(x - z_j)^2}\right]\right| \cdot \left|\frac{f(x)}{f'(x)}\right| \\
&\leq \frac{1}{2} \cdot \left[\frac{1}{m} \cdot \left(\sum_{j=1}^{n} \left|\frac{1}{x - z_j}\right|\right)^2 + \sum_{j=1}^{n} \frac{1}{|x - z_j|^2}\right] \cdot \left|\frac{f(x)}{f'(x)}\right| \\
&\leq \frac{1}{2} \cdot \left(\frac{n^2}{m} \cdot \frac{1}{|x - z|^2} + \frac{n}{|x - z|^2}\right) \cdot \left|\frac{f(x)}{f'(x)}\right| \\
&= \frac{n^2 + mn}{2m} \cdot \left|\frac{f(x)}{f'(x)}\right| \cdot \frac{1}{|x - z|^2}\ ,
\end{aligned}$$

$$|x-z|^2 \leq \frac{n^2+nm}{2m} \cdot \left|\frac{f(x)}{f'(x)}\right| \cdot \frac{1}{|F_m(x)|} \,.$$

Therefore,

$$|x-z|^2 \leq \frac{n^2+nm}{2m} \cdot \left|\frac{f(x)}{f'(x)}\right| \cdot |H_m(x)-x| \,.$$

□

It is not clear in general which one is the best among the three inequalities in (2.18), (2.19) and (2.21). For the following two examples, we use $m = 1$ in (2.19) and (2.21). For

$$f(x) = x^6 - 4x^5 - 5x^4 + 190x^3 - 666x^2 + 944x - 600$$

with degree 6 and zeros $-6, 2, 1 \pm i, 3 \pm 4i$, let $x_0 = 0$. Then, $f(x_0) = -600$, $f'(x_0) = 944$, $|L_1^-(x_0) - x_0| = |L_1^+(x_0) - x_0| = 2.238$ and $|H_1(x_0) - x_0| = 1.152$. By Kahan's inequality (2.19),

$$LB = \sqrt{\frac{n}{m}} |L_m^{\pm}(x_0) - x_0| \doteq \sqrt{6}(2.238) \doteq 5.482.$$

By inequality (2.21),

$$\begin{aligned} HB &= \sqrt{\frac{n^2+nm}{2m}\left|\frac{f(x_0)}{f'(x_0)}\right| |H_m(x_0)-x_0|} \\ &= \sqrt{\frac{6^2+6}{2}\left|\frac{-600}{944}\right|(1.152)} \\ &\doteq 3.921. \end{aligned}$$

By Laguerre's inequality (2.18),

$$NB = n\left|\frac{f(x_0)}{f'(x_0)}\right| = 6\left|\frac{-600}{944}\right| = 3.814.$$

Thus, for this example, we have

$$LB > HB > NB.$$

For

$$f(x) = x^6 - 6x^5 + 50x^3 - 45x^2 - 108x + 108$$

with degree 6 and zeros $1, -2, -2, 3, 3, 3$, let $x_0 = 0$. We have $f(x_0) = 108$, $f'(x_0) = -108$, $|L_1^-(x_0) - x_0| > |L_1^+(x_0) - x_0| = 0.743$ and $|H_1(x_0) - x_0| = 0.706$. By Kahan's inequality (2.19),

$$LB = \sqrt{\frac{n}{m}}|L_m^+(x_0) - x_0| \doteq \sqrt{6}(0.743) \doteq 1.820.$$

Inequality (2.21) gives

$$\begin{aligned} HB &= \sqrt{\frac{n^2+nm}{2m}\left|\frac{f(x_0)}{f'(x_0)}\right||H_m(x_0)-x_0|} \\ &= \sqrt{\frac{6^2+6}{2}\left|\frac{108}{-108}\right|(0.706)} \\ &\doteq 3.851. \end{aligned}$$

and by Laguerre's inequality (2.18),

$$NB = n\left|\frac{f(x_0)}{f'(x_0)}\right| = 6\left|\frac{108}{-108}\right| = 6.$$

Thus, for this example, we have

$$LB < HB < NB.$$

2.5 The two-sided modified Halley iteration

The modified Laguerre iteration defined by Li and Zeng [36] is

$$x_{k+1} = L_m^{\pm}(x_k), \quad k = 0, 1, 2, \ldots$$

where $L_m^{\pm}(x)$ is the modified Laguerre iteration function in (2.20). This iteration is two-sided in the sense that for a given x-value two corresponding approximation values $L_m^+(x)$ and $L_m^-(x)$ are obtained. Similarly, for the modified Halley iteration (2.9), we construct the **Two-sided Modified Halley Iteration** as follows:

$$x_{k+1} = x_k \pm \frac{2m|f(x_k)f'(x_k)|}{(1+m)[f'(x_k)]^2 - mf(x_k)f''(x_k)} \equiv H_m^{\pm}(x_k) \tag{2.22}$$

where

$$H_m^{\pm}(x) = x \pm \frac{2m|f(x)f'(x)|}{(1+m)[f'(x)]^2 - mf(x)f''(x)}.$$

In the rest of this section, we proceed to show the non-overshoot properties of this two-sided iteration.

Theorem 2.5.1 *Let $f(x)$ be a polynomial with real zeros*

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n \ .$$

Then for any $x \in (\lambda_i, \lambda_{i+1})$ with $+\infty > |\frac{f(x)}{f'(x)}| > 0$ and any positive integer m , we have

$$\lambda_i < x < H_1^+(x) < H_2^+(x) < \cdots < H_m^+(x) < \lambda_{i+m} \tag{2.23}$$

and

$$\lambda_{i-m+1} < H_m^-(x) < H_{m-1}^-(x) < \cdots < H_1^-(x) < x < \lambda_{i+1} \tag{2.24}$$

with the convention $\lambda_i = -\infty$ for $i \leq 0$ and $\lambda_i = +\infty$ for $i \geq n+1$.

Proof Since

$$H_m^\pm(x) = x \pm \frac{|f(x)f'(x)|}{\left(\frac{m+1}{2m}\right)(f'(x))^2 - \frac{1}{2}f(x)f''(x)} \ ,$$

we have

$$\frac{\partial H_m^+(x)}{\partial m} = -\frac{\partial H_m^-(x)}{\partial m} = \frac{|f(x)f'(x)|(f'(x))^2}{2m^2\left(\left(\frac{m+1}{2m}\right)(f'(x))^2 - \frac{1}{2}f(x)f''(x)\right)^2} > 0 \ .$$

Rewrite

$$H_m^\pm(x) = x \pm \frac{1}{F_m(x)}$$

where

$$\begin{aligned} F_m(x) &= \frac{(1+m)(f'(x))^2 - mf(x)f''(x)}{2m|f(x)f'(x)|} \\ &= \frac{(f'(x))^2 + m[(f'(x))^2 - f(x)f''(x)]}{2m(f(x))^2} \cdot \left|\frac{f(x)}{f'(x)}\right| \\ &= \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right|^2 + \frac{(f')^2 - f\cdot f''}{(f)^2}\right) \cdot \left|\frac{f(x)}{f'(x)}\right| \\ &= \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right| + \left|\frac{f(x)}{f'(x)}\right| \cdot \sum_{j=1}^{n} \frac{1}{(x-\lambda_j)^2}\right) \ . \end{aligned}$$

Since

$$F_m(x) > \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right| + \left|\frac{f(x)}{f'(x)}\right| \cdot \sum_{j=i+1}^{i+m} \frac{1}{(x-\lambda_j)^2}\right)$$

$$
\begin{aligned}
&> \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right| + \left|\frac{f(x)}{f'(x)}\right| \cdot \frac{m}{(x-\lambda_{i+m})^2}\right) \\
&\geq \frac{1}{|x-\lambda_{i+m}|} \qquad (\text{ since } \frac{1}{2}(a+b) \geq \sqrt{ab}\ ) \\
&= \frac{1}{\lambda_{i+m}-x} > 0\ ,
\end{aligned}
$$

we have

$$
x < H_m^+(x) = x + \frac{1}{F_m(x)} < x + \lambda_{i+m} - x = \lambda_{i+m}\ .
$$

Thus, (2.23) is true by $\frac{\partial H_m^+(x)}{\partial m} > 0$.

Since

$$
\begin{aligned}
F_m(x) &> \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right| + \left|\frac{f(x)}{f'(x)}\right| \cdot \sum_{j=i-m+1}^{i} \frac{1}{(x-\lambda_j)^2}\right) \\
&> \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x)}{f(x)}\right| + \left|\frac{f(x)}{f'(x)}\right| \cdot \frac{m}{(x-\lambda_{i-m+1})^2}\right) \\
&\geq \frac{1}{|x-\lambda_{i-m+1}|} \qquad (\text{ since } \frac{1}{2}(a+b) \geq \sqrt{ab}\ ) \\
&= \frac{1}{x-\lambda_{i-m+1}} > 0\ ,
\end{aligned}
$$

we have

$$
x > H_m^-(x) = x - \frac{1}{F_m(x)} > x - (x - \lambda_{i-m+1}) = \lambda_{i-m+1}\ .
$$

Thus, (2.24) is true since $\frac{\partial H_m^-(x)}{\partial m} < 0$. □

The following corollary is a direct consequence of Theorem 2.5.1 .

Corollary 2.5.1 *Let*

$$
x_{k+1}^{\pm} = H_1^{\pm}(x_k), \quad k = 0, 1, 2, \ldots .
$$

Then, for any $x_0 \in (\lambda_i, \lambda_{i+1})$ *with*

$$
+\infty > \left|\frac{f(x_k^{\pm})}{f'(x_k^{\pm})}\right| > 0, \quad k = 0, 1, 2, \ldots,
$$

we have

$$
\lambda_i \overset{cubically}{\longleftarrow} x_k^- < \cdots < x_2^- < x_1^- < x_0 < x_1^+ < x_2^+ < \cdots < x_k^+ \overset{cubically}{\longrightarrow} \lambda_{i+1}\ .
$$

Theorem 2.5.2 *Suppose*

$$f(x) = c(x-\lambda)^m \cdot \prod_{j=1}^{n-m} (x-\lambda_j)$$

where c is a constant. Let

$$x_{k+1}^{\pm} = H_m^{\pm}(x_k^{\pm}), \quad k = 0, 1, 2, \ldots .$$

If $x_0 < \lambda$ *and* $0 < \left|\frac{f(x_k^+)}{f'(x_k^+)}\right| < +\infty$ *for* $k = 0, 1, 2, \ldots,$ *then*

$$x_0 < x_1^+ < x_2^+ < \cdots < x_k^+ \overset{cubically}{\longrightarrow} \lambda \, . \tag{2.25}$$

If $\lambda < x_0$ *and* $0 < \left|\frac{f(x_k^-)}{f'(x_k^-)}\right| < +\infty$ *for* $k = 0, 1, 2, \ldots,$ *then*

$$\lambda \overset{cubically}{\longleftarrow} x_k^- < x_{k-1}^- < \cdots < x_1^- < x_0 \, . \tag{2.26}$$

Proof It follows from the proof of Theorem 2.5.1,

$$\begin{aligned} F_m(x_0) &= \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x_0)}{f(x_0)}\right| + \left|\frac{f(x_0)}{f'(x_0)}\right| \cdot \frac{(f'(x_0))^2 - f(x_0)f''(x_0)}{(f(x_0))^2}\right) \\ &= \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x_0)}{f(x_0)}\right| + \left|\frac{f(x_0)}{f'(x_0)}\right| \left(\frac{m}{(x_0-\lambda)^2} + \sum_{j=1}^{n-m} \frac{1}{(x_0-\lambda_j)^2}\right)\right) \\ &> \frac{1}{2}\left(\frac{1}{m}\left|\frac{f'(x_0)}{f(x_0)}\right| + \left|\frac{f(x_0)}{f'(x_0)}\right| \cdot \frac{m}{(x_0-\lambda)^2}\right) \\ &\geq \frac{1}{|x_0-\lambda|} \, . \end{aligned}$$

If $x_0 < \lambda$, then

$$F_m(x_0) > \frac{1}{\lambda - x_0} > 0 \, .$$

So

$$x_0 < x_1^+ = H_m^+(x_0) = x_0 + \frac{1}{F_m(x_0)} < x_0 + (\lambda - x_0) = \lambda \, .$$

If $x_0 > \lambda$, then

$$F_m(x_0) > \frac{1}{x_0 - \lambda} > 0 \, .$$

So

$$x_0 > x_1^- = H_m^-(x_0) = x_0 - \frac{1}{F_m(x_0)} > x_0 - (x_0 - \lambda) = \lambda \, .$$

□

2.6 Halley's iteration in real or complex Banach space

One of the basic results in numerical analysis is the Newton-Kantorovich Theorem which provides a sufficient condition that guarantees the local convergence of the Newton iteration. This classical result can be found in Schröder [47], Jankó [28], Šafiev [45], Kantorovich and Akilov [31], Döring [14], Ortega and Rheinboldt [38] and Ostrowski [39]. More recent work in this direction includes Gragg and Tapia [24], Rall [43], Miel [37] and Traub and Wozniakowski [52]. In this section we first derive Halley's iteration in Banach space followed by a Newton-Kantorovich type convergence theorem for Halley's iteration in Banach space. We show that the convergence rate for Halley's iteration in Banach space is cubic, along with an optimal error estimation for the iteration.

2.6.1 Derivation for Halley's iteration in Banach space

Let X and Y be Banach spaces and $D \subset X$ be a convex subset of X. Let $F : D \subset X \to Y$ be a twice Frechet-differentiable nonlinear operator on D. If $F(X^*) = 0$, then

$$\begin{aligned} 0 &= F(X^*) \doteq F(X_k) + F'(x_k)(X^* - X_k) + \frac{1}{2}F''(X_k)(X^* - X_k)(X^* - X_k) \\ &= F(X_k) + F'(X_k)\left\{I + \frac{1}{2}[F'(X_k)]^{-1}F''(X_k)(X^* - X_k)\right\}(X^* - X_k). \end{aligned}$$

So

$$X^* \doteq X_k - \left\{I + \frac{1}{2}[F'(X_k)]^{-1}F''(X_k)(X^* - X_k)\right\}^{-1}[F'(X_k)]^{-1}F(X_k).$$

Now replace $X^* - X_k$ by Newton's correction $-[F'(X_k)]^{-1}F(X_k)$ on the right-hand side of the above equation and call the resulting right-hand side X_{k+1}, we have

$$X_{k+1} = X_k - \left\{I - \frac{1}{2}[F'(X_k)]^{-1}F''(X_k)[F'(X_k)]^{-1}F(X_k)\right\}^{-1}[F'(X_k)]^{-1}F(X_k) \quad (2.27)$$

which we call **Halley's Iteration in Banach Space** . let

$$\Gamma_k = [F'(X_k)]^{-1}$$

and

$$\Theta_k = \left\{ I - \frac{1}{2}\Gamma_k F''(X_k)\Gamma_k F(X_k) \right\}^{-1},$$

then the above Halley iteration in Banach space can be rewritten as

$$X_{k+1} = X_k - \Theta_k \Gamma_k F(X_k), \quad k = 0, 1, 2, \ldots . \tag{2.28}$$

2.6.2 Convergence of Halley's iteration in Banach space

Let

$$S(X_0, r) = \{X \mid ||X - X_0|| < r\}$$

and

$$\overline{S(X_0, r)} = \{X \mid ||X - X_0|| \le r\}.$$

The following convergence theorem for Halley's iteration in Banach space is proved by Šafiev [44].

Theorem (Šafiev) Let $F : D \subset X \to Y$ be a three times Frechet-differentiable nonlinear operator with $||F''(X)|| \le M$ and $||F'''(X)|| \le N$ on the convex set D. Suppose that $X_0 \in D$ satisfies the following conditions:

- $||\Gamma_0|| \le B_0$ and $||F(X_0)|| \le \delta_0$;
- $h_0 = MB_0^2\delta_0 \le \frac{1}{2+\gamma_0}$ with $\gamma_0 = NB_0^{-1}M^{-2}$;
- $S(X_0, t^*) \subset D$ where t^* is the smallest positive root of the following equation

$$\varphi(t) \equiv \frac{1}{6}Nt^3 + \frac{1}{2}Mt^2 - B_0^{-1}t + \delta_0 = 0 .$$

Then, Halley's iterations $X_{k+1} = X_k - \Theta_k\Gamma_k F(X_k)$ as well as its limit $X^* = \lim_{k\to\infty} X_k$ exist with $F(X^*) = 0$, and

$$||X^* - X_k|| \le t^* - t_k$$

where $\{t_k\}$ are determined by applying one-dimensional Halley's iteration (2.2) to the function $\varphi(t)$ with $t_0 = 0$ and limit t^*.

The above Theorem (Šafiev) does not give an explicit error bound for $||X^* - X_k||$. Therefore, in the following, we give another convergence theorem for Halley's iteration in Banach space similar to the Newton-Kantorovich Theorem with the optimal error bounds given by Gragg and Tapia [24].

Theorem 2.6.1 *Let $F : D \subset X \to Y$ be a three times Frechet-differentiable nonlinear operator with $||F''(X)|| \leq M$ and $||F'''(X)|| \leq N$ on the convex set D. Suppose that $X_0 \in D$ satisfies the following conditions:*

- $||\Gamma_0|| \leq B_0$ *and* $||F(X_0)|| \leq B_0^{-1}\eta_0$;
- $h_0 = KB_0\eta_0 \leq \frac{1}{2}$ *with* $K = \sqrt{M^2 + \frac{2}{3}\frac{N}{B_0(1-\frac{1}{2}MB_0\eta_0)}}$;
- $S(X_0, (1+\theta_0)\eta_0) \subset D$ *where* $\theta_0 = \frac{1-\sqrt{1-2h_0}}{1+\sqrt{1-2h_0}}$.

Then

1. *The Halley iteration* $X_{k+1} = X_k - \Theta_k\Gamma_k F(X_k)$ *exists and* $X_k \in S(X_0, (1+\theta_0)\eta_0), \quad k = 0, 1, 2, \ldots$;
2. $X^* = \lim_{k\to\infty} X_k$ *exists,* $X^* \in \overline{S(X_0, (1+\theta_0)\eta_0)}$ *and* $F(X^*) = 0$;
3. *The optimal error bound is*
$$||X^* - X_k|| \leq \frac{(1+\theta_0)\eta_0\theta_0^{3^k-1}}{\sum_{i=0}^{3^k-1}\theta_0^i}, \quad k = 1, 2, \ldots. \tag{2.29}$$
Moreover, if $h_0 < \frac{1}{2}$, *i.e.,* $\theta_0 < 1$, *then*
$$||X^* - X_k|| \leq \frac{\theta_0^{3^k}}{1-\theta_0^{3^k}}(\theta_0^{-1} - \theta_0)\eta_0.$$
So the order of convergence for Halley's iteration in Banach space is cubic.
4. *If* $\overline{S(X_0, (1+\theta_0^{-1})\eta_0)} \subset D$, *then* X^* *is the unique solution of* $F(X) = 0$ *in the set*
$$\overline{S(X_0, (1+\theta_0)\eta_0)} \bigcup S(X_0, (1+\theta_0^{-1})\eta_0).$$

Remark 2.6.1 *The error bound (2.29) in Theorem 2.6.1 is optimal in the sense that one may construct a function with initial value X_0 that satisfies the conditions of Theorem 2.6.1 for which the error bound (2.29) holds with equality.*

Remark 2.6.2 *When X and Y are both one-dimensional, a similar convergence theorem for Halley's iteration (2.2) is proved by Zheng [58].*

2.6.3 The proof of the convergence theorem for Halley's iteration in Banach space

To prove Theorem 2.6.1 we need the following definition and lemmas.

Definition 2.6.1 *Let $\{X_k\}$ be any sequence in X. A sequence $\{t_k\} \subset [0, \infty) \subset R^1$ for which*

$$||X_{k+1} - X_k|| \leq t_{k+1} - t_k, \quad k = 0, 1, 2, \ldots,$$

holds is a majorizing sequence for $\{X_k\}$.

Lemma 2.6.1 *Let $\{t_k\} \subset [0, \infty)$ be a majorizing sequence for $\{X_k\} \subset X$, and suppose that $\lim_{k\to\infty} t_k = t^* < \infty$ exists. Then $X^* = \lim_{k\to\infty} X_k$ exists and*

$$||X^* - X_k|| \leq t^* - t_k, \quad k = 0, 1, 2, \ldots. \tag{2.30}$$

Proof Since

$$||X_{k+m} - X_k|| \leq \sum_{j=k}^{k+m-1} ||X_{j+1} - X_j|| \tag{2.31}$$

$$\leq \sum_{j=k}^{k+m-1} (t_{j+1} - t_j) = t_{k+m} - t_k\ , \tag{2.32}$$

$\{X_k\}$ is a Cauchy sequence and the error estimate (2.30) follows from (2.32) as $m \to \infty$. □

Lemma 2.6.2 (Neumann Lemma) *If $A : X \to Y$ is a bounded linear operator and $||A|| < 1$, then $(I - A)^{-1}$ exists , $(I - A)^{-1} = \lim_{k\to\infty} \sum_{i=0}^{k} A^i$ and*

$$||(I - A)^{-1}|| \leq \sum_{i=0}^{\infty} ||A||^i = \frac{1}{1 - ||A||}.$$

Lemma 2.6.3 *If $F : D \subset X \to Y$ has a second Frechet-derivative at each point of the convex set D, then for any $X, Y \in D$,*

$$||F'(Y) - F'(X)|| \leq ||Y - X|| \sup_{0\leq t\leq 1} ||F''(X + t(Y - X))||.$$

The proofs for Lemma 2.6.2 and Lemma 2.6.3 can be found in Dieudonné [11] and Lang [34].

Proof of Theorem 2.6.1 We divide the proof into three parts:

Part 1. *The error bound in (2.29) is optimal.*

Let

$$\phi(t) = \frac{K}{2}t^2 - B_0^{-1}t + B_0^{-1}\eta_0$$

where K, B_0, η_0 are the constants as in Theorem 2.6.1. Then

$$\phi(t) = \frac{K}{2}(t - r_1)(t - r_2)$$

and

$$\phi'(t) = \frac{K}{2}[(t - r_1) + (t - r_2)] \text{ and } \phi''(t) = K$$

where $r_1 = (1 + \theta_0)\eta_0$ and $r_2 = (1 + \theta_0^{-1})\eta_0$ with $0 < r_1 \leq r_2$. Let $t_0 = 0$ and apply the Halley iteration to the scalar polynomial $\phi(t)$. We have

$$t_{k+1} = t_k - \frac{\phi(t_k)}{\phi'(t_k) - \frac{1}{2}\frac{\phi(t_k)\phi''(t_k)}{\phi'(t_k)}}, \quad k = 0, 1, 2, \ldots. \tag{2.33}$$

By Corollary 2.3.1, the sequence $\{t_k\}$ converges to r_1 monotonically, and by (2.33),

$$\begin{aligned} t_{k+1} &= t_k - \frac{(t_k - r_1)(t_k - r_2)}{t_k - r_1 + t_k - r_2 - \frac{(t_k - r_1)(t_k - r_2)}{t_k - r_1 + t_k - r_2}} \\ &= t_k - \frac{(t_k - r_1)(t_k - r_2)[(t_k - r_1) + (t_k - r_2)]}{(t_k - r_1)^2 + (t_k - r_2)^2 + (t_k - r_1)(t_k - r_2)}. \end{aligned}$$

So

$$t_{k+1} - r_1 = \frac{(t_k - r_1)^3}{(t_k - r_1)^2 + (t_k - r_2)^2 + (t_k - r_1)(t_k - r_2)}. \tag{2.34}$$

Similarly,

$$t_{k+1} - r_2 = \frac{(t_k - r_2)^3}{(t_k - r_1)^2 + (t_k - r_2)^2 + (t_k - r_1)(t_k - r_2)}. \tag{2.35}$$

Let $\theta_k = \frac{r_1 - t_k}{r_2 - t_k}$, then by (2.34) and (2.35),

$$\theta_{k+1} = \theta_k^3, \quad k = 0, 1, 2, \ldots.$$

Thus

$$\frac{r_1 - t_k}{r_2 - r_1 + r_1 - t_k} = \theta_k = \theta_{k-1}^3 = \theta_{k-2}^{3^2} = \cdots = \theta_0^{3^k}.$$

So, if $h_0 < \frac{1}{2}$ (i.e., $\theta_0 < 1$), then

$$r_1 - t_k = \frac{\theta_0^{3^k}}{1-\theta_0^{3^k}}(r_2 - r_1) = \frac{\theta_0^{3^k}}{1-\theta_0^{3^k}}(\theta_0^{-1} - \theta_0)\eta_0 \tag{2.36}$$

$$= \frac{(1+\theta_0)\eta_0\theta_0^{3^k-1}}{\sum_{i=0}^{3^k-1}\theta_0^i}. \tag{2.37}$$

If $h_0 = \frac{1}{2}$ (i.e., $\theta_0 = 1$), then $r_2 = r_1$. By (2.34),

$$r_1 - t_{k+1} = \frac{1}{3}(r_1 - t_k) = \cdots = \frac{1}{3^k}(r_1 - t_0) = \frac{(1+\theta_0)\eta_0}{3^k}.$$

So the equation (2.37) is still true for $\theta_0 = 1$ and holds for all $h_0 \leq \frac{1}{2}$. Therefore, the bound (2.29) in Theorem 2.6.1 is optimal.

Part 2. *The sequence $\{t_k\}$ in Part 1 is a majorizing sequence for $\{X_k\}$ generated by Halley's iteration (2.28).*

We shall use mathematical induction to prove:

$$||F(X_k)|| \leq \phi(t_k); \tag{2.38}$$

$$||\Gamma_k|| \leq -\phi'(t_k)^{-1}; \tag{2.39}$$

$$||\Theta_k|| \leq \frac{1}{1 - \frac{M}{2}\frac{\phi(t_k)}{\phi'(t_k)^2}}; \tag{2.40}$$

$$||X_{k+1} - X_k|| \leq t_{k+1} - t_k \tag{2.41}$$

where $k = 0, 1, 2, \ldots$, and $\phi(t)$ is the same scalar function as defined in Part 1.

When $k = 0$,

$$||F(x_0)|| \leq B_0^{-1}\eta_0 = \phi(t_0)$$

and

$$||\Gamma_0|| \leq B_0 = -\phi'(t_0)^{-1}.$$

Since

$$\left|\left|\frac{1}{2}\Gamma_0 F''(X_0)\Gamma_0 F(X_0)\right|\right| \leq \frac{1}{2}M\frac{\phi(t_0)}{\phi'(t_0)^2} = \frac{1}{2}MB_0\eta_0 \leq \frac{1}{2}KB_0\eta_0 < 1,$$

we have,

$$||\Theta_0|| = \left|\left|[I - \frac{1}{2}\Gamma_0 F''(X_0)\Gamma_0 F(X_0)]^{-1}\right|\right| \leq \frac{1}{1 - \frac{1}{2}M\frac{\phi(t_0)}{\phi'(t_0)^2}}.$$

So,

$$\begin{aligned} ||X_1 - X_0|| &= ||\Theta_0\Gamma_0 F(X_0)|| \leq ||\Theta_0|| \cdot ||\Gamma_0|| \cdot ||F(X_0)|| \\ &\leq \frac{-\frac{\phi(t_0)}{\phi'(t_0)}}{1 - \frac{1}{2}M\frac{\phi(t_0)}{\phi'(t_0)^2}} \\ &\leq \frac{-\frac{\phi(t_0)}{\phi'(t_0)}}{1 - \frac{1}{2}K\frac{\phi(t_0)}{\phi'(t_0)^2}} = t_1 - t_0. \end{aligned}$$

Thus, (2.38)–(2.41) are true for $k = 0$.

Now suppose that (2.38)–(2.41) are true for all $k \leq n$. Then,

$$\begin{aligned} ||X_{n+1} - X_0|| &= ||\sum_{k=0}^{n}(X_{k+1} - X_k)|| \\ &\leq \sum_{k=0}^{n}(t_{k+1} - t_k) = t_{n+1} - t_0 = t_{n+1} \leq r_1. \end{aligned}$$

So,

$$X_{n+1} \in S(X_0, r_1).$$

By Lemma 2.6.3,

$$\begin{aligned} ||[F'(X_0)]^{-1}[F'(X_0) - F'(X_{n+1})]|| &\leq B_0 M ||X_0 - X_{n+1}|| \\ &\leq B_0 M (t_{n+1} - t_0) \\ &= B_0 K t_{n+1} \leq B_0 K r_1 \\ &= B_0 K \frac{\frac{1}{B_0} - \sqrt{\frac{1}{B_0^2} - \frac{2K\eta_0}{B_0}}}{K} < 1. \end{aligned}$$

On the other hand, by Lemma 2.6.2 and

$$\begin{aligned} F'(X_{n+1}) &= F'(X_0) - [F'(X_0) - F'(X_{n+1})] \\ &= F'(X_0)\{I - [F'(X_0)]^{-1}[F'(X_0) - F'(X_{n+1})]\}, \end{aligned}$$

we have,

$$\begin{aligned} ||\Gamma_{n+1}|| &= ||[F'(X_{n+1})]^{-1}|| \\ &= ||\{I - [F'(X_0)]^{-1}[F'(X_0) - F'(X_{n+1})]\}^{-1}[F'(X_0)]^{-1}|| \\ &\leq \frac{B_0}{1 - B_0 K t_{n+1}} = \frac{1}{\frac{1}{B_0} - K t_{n+1}} = -\phi'(t_{n+1})^{-1}. \end{aligned}$$

So (2.39) is true for $k = n+1$.

For (2.38), let

$$\tilde{F_n} = F(X_n) + F'(X_n)(X - X_n) - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X - X_n). \quad (2.42)$$

Then,

$$\begin{aligned}\tilde{F_n} &= F(X_n) + F'(X_n)(X_{n+1} - X_n) - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X_{n+1} - X_n)\\ &\quad + F'(X_n)(X - X_{n+1}) - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X - X_{n+1})\\ &= F'(X_n)(X - X_{n+1}) - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X - X_{n+1})\\ &= F'(X_n)\{I - \frac{1}{2}[F'(X_n)]^{-1}F''(X_n)[F'(X_n)]^{-1}F(X_n)\}(X - X_{n+1}).\end{aligned}$$

So,

$$X - X_{n+1} = \Theta_n \Gamma_n \tilde{F_n}. \quad (2.43)$$

where $\tilde{F_n}$ can be rewritten as

$$\begin{aligned}\tilde{F_n} &= F(X_n) + F'(X_n)(X - X_n) + \frac{1}{2}F''(X_n)(X - X_n)(X - X_n)\\ &\quad - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X - X_n) - \frac{1}{2}F''(X_n)(X - X_n)(X - X_n)\\ &= F(X_n) + F'(X_n)(X - X_n) + \frac{1}{2}F''(X_n)(X - X_n)(X - X_n)\\ &\quad - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}\{F(X_n) + [F'(X_n)](X - X_n)\}(X - X_n). \quad (2.44)\end{aligned}$$

Now let $X = X_{n+1}$ in (2.43) and (2.44), then

$$\begin{aligned}0 = \tilde{F_n} &= F(X_n) + F'(X_n)(X_{n+1} - X_n) - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X_{n+1} - X_n)\\ &= F(X_n) + F'(X_n)(X_{n+1} - X_n) + \frac{1}{2}F''(X_n)(X_{n+1} - X_n)(X_{n+1} - X_n)\\ &\quad - \frac{1}{2}F''(X_n)[F'(X_n)]^{-1}\{F(X_n) + [F'(X_n)](X_{n+1} - X_n)\}(X_{n+1} - X_n)\\ &= F(X_n) + F'(X_n)(X_{n+1} - X_n) + \frac{1}{2}F''(X_n)(X_{n+1} - X_n)(X_{n+1} - X_n)\\ &\quad - \frac{1}{4}F''(X_n)[F'(X_n)]^{-1}F''(X_n)[F'(X_n)]^{-1}F(X_n)(X_{n+1} - X_n)(X_{n+1} - X_n).\end{aligned}$$

Thus,

$$F(X_{n+1}) = \{F(X_{n+1}) - F(X_n) - F'(X_n)(X_{n+1} - X_n)$$

$$-\frac{1}{2}F''(X_n)(X_{n+1}-X_n)(X_{n+1}-X_n)\}$$
$$+\{\frac{1}{4}F''(X_n)[F'(X_n)]^{-1}F''(X_n)[F'(X_n)]^{-1}$$
$$F(X_n)(X_{n+1}-X_n)(X_{n+1}-X_n)\}. \tag{2.45}$$

Similarly, since $\phi(t)$ is a quadratic polynomial, we have

$$\phi(t_{n+1}) = \frac{K^2}{4}\cdot\frac{\phi(t_n)}{(\phi'(t_n))^2}(t_{n+1}-t_n)^2 \tag{2.46}$$

By (2.45) and the induction assumption, we have

$$\begin{aligned}
||F(X_{n+1})|| &\le \frac{N}{6}||X_{n+1}-X_n||^3 + \frac{M^2}{4}||[F'(X_n)]^{-1}||^2||F(X_n)||||X_{n+1}-X_n||^2 \\
&\le \left\{\frac{N}{6}||X_{n+1}-X_n|| + \frac{M^2}{4}\frac{\phi(t_n)}{(\phi'(t_n))^2}\right\}(t_{n+1}-t_n)^2 \\
&= \left\{\frac{N}{6}||\Theta_n\Gamma_nF(X_n)|| + \frac{M^2}{4}\frac{\phi(t_n)}{(\phi'(t_n))^2}\right\}(t_{n+1}-t_n)^2 \\
&\le \left\{\frac{N}{6}\frac{-\frac{\phi(t_n)}{\phi'(t_n)}}{1-\frac{M}{2}\frac{\phi(t_n)}{(\phi'(t_n))^2}} + \frac{M^2}{4}\frac{\phi(t_n)}{(\phi'(t_n))^2}\right\}(t_{n+1}-t_n)^2 \\
&= \left\{\frac{N}{6}\frac{-\phi'(t_n)}{1-\frac{M}{2}\frac{\phi(t_n)}{(\phi'(t_n))^2}} + \frac{M^2}{4}\right\}\frac{\phi(t_n)}{(\phi'(t_n))^2}(t_{n+1}-t_n)^2.
\end{aligned}$$

Now let

$$g(t) = \frac{\phi(t)}{(\phi'(t))^2}.$$

When $r_1 = r_2$, we have $g(t) = \frac{1}{2K}$. When $r_1 \ne r_2$, then

$$g'(t) = \frac{(r_2-r_1)^2}{\frac{K}{2}[2t-r_1-r_2]^3} < 0, \quad \forall t \in [0, r_1].$$

So,

$$\frac{\phi(t_n)}{(\phi'(t_n))^2} \le \frac{\phi(t_0)}{(\phi'(t_0))^2}.$$

Since $-\phi'(t_n) \le -\phi'(t_0)$,

$$\frac{-\phi'(t_n)}{1-\frac{M}{2}\frac{\phi(t_n)}{(\phi'(t_n))^2}} \le \frac{-\phi'(t_0)}{1-\frac{M}{2}\frac{\phi(t_0)}{(\phi'(t_0))^2}} = \frac{1}{B_0\left(1-\frac{M}{2}B_0\eta_0\right)}.$$

Thus,

$$\begin{aligned}
||F(X_{n+1})|| &\le \left[M^2 + \frac{2}{3}\frac{N}{B_0\left(1-\frac{M}{2}B_0\eta_0\right)}\right]\frac{1}{4}\frac{\phi(t_n)}{(\phi'(t_n))^2}(t_{n+1}-t_n)^2 \\
&= \frac{K^2}{4}\frac{\phi(t_n)}{(\phi'(t_n))^2}(t_{n+1}-t_n)^2 = \phi(t_{n+1}).
\end{aligned}$$

So (2.38) is true for $k = n+1$. Since

$$\begin{aligned}
\left\|\frac{1}{2}\Gamma_{n+1}F''(X_{n+1})\Gamma_{n+1}F(X_{n+1})\right\| &\le \frac{M}{2}\frac{\phi(t_{n+1})}{(\phi'(t_{n+1}))^2} \\
&\le \frac{M}{2}\frac{\phi(t_0)}{(\phi'(t_0))^2} \\
&= \frac{1}{2}MB_0\eta_0 < 1,
\end{aligned}$$

$$\begin{aligned}
||\Theta_{n+1}|| &= \left\|\left\{I - \frac{1}{2}\Gamma_{n+1}F''(X_{n+1})\Gamma_{n+1}F(X_{n+1})\right\}^{-1}\right\| \\
&\le \frac{1}{1 - \frac{M}{2}\frac{\phi(t_{n+1})}{\phi'(t_{n+1})^2}}
\end{aligned}$$

and

$$\begin{aligned}
||X_{n+2} - X_{n+1}|| &= ||\Theta_{n+1}\Gamma_{n+1}F(X_{n+1})|| \\
&\le \frac{-\frac{\phi(t_{n+1})}{\phi'(t_{n+1})}}{1 - \frac{M}{2}\frac{\phi(t_{n+1})}{\phi'(t_{n+1})^2}} \\
&\le \frac{-\frac{\phi(t_{n+1})}{\phi'(t_{n+1})}}{1 - \frac{K}{2}\frac{\phi(t_{n+1})}{\phi'(t_{n+1})^2}} \\
&= t_{n+2} - t_{n+1}.
\end{aligned}$$

Thus, (2.40) and (2.41) are true for $k = n+1$. Therefore, (2.38)–(2.41) hold for all $k \ge 0$.

Consequently, by Lemma 2.6.1,

$$\lim_{k\to\infty} X_k = X^*$$

exists with

$$\begin{aligned}
||X^* - X_k|| &\le r_1 - t_k \\
&= (1+\theta_0)\eta_0 \frac{\theta_0^{3^k-1}}{\sum_{i=0}^{3^k-1}\theta_0^i}, \quad k = 0, 1, 2, \ldots,
\end{aligned}$$

and $X^* \in \overline{S(X_0, r_1)}$.

Part 3. *X^* is the unique solution of $F(X) = 0$ in*

$$\overline{S(X_0, (1+\theta_0)\eta_0)} \bigcup S(X_0, (1+\theta_0^{-1})\eta_0).$$

If $\bar{X} \in \overline{S(X_0, r_1)} \bigcup S(X_0, r_2)$ is another solution of $F(X) = 0$, then letting $X = \bar{X}$ in (2.43) and (2.44), we have

$$\bar{X} - X_{n+1} = \Theta_n \Gamma_n \tilde{F}_n$$

where

$$\begin{aligned} \tilde{F}_n &= \left\{ F(X_n) + F'(X_n)(\bar{X} - X_n) + \frac{1}{2} F''(X_n)(\bar{X} - X_n)(\bar{X} - X_n) - F(\bar{X}) \right\} \\ &+ \frac{1}{2} F''(X_n)[F'(X_n)]^{-1}\{F(\bar{X}) - F(X_n) - F'(X_n)(\bar{X} - X_n)\}(\bar{X} - X_n). \end{aligned} \tag{2.47}$$

Similarly,

$$r_2 - t_{n+1} = \frac{\frac{K^2}{4}(r_2 - t_n)^3 / (\phi'(t_n))^2}{1 - \frac{K}{2}\phi(t_n) / (\phi'(t_n))^2}. \tag{2.48}$$

By (2.47),

$$\begin{aligned} ||\tilde{F}_n|| &\le \left[\frac{N}{6} + \frac{M^2}{4} ||[F'(X_n)]^{-1}||\right] ||\bar{X} - X_n||^3 \\ &\le \left[\frac{N}{6} - \frac{M^2}{4\phi'(t_n)}\right] ||\bar{X} - X_n||^3 \\ &= \left[-\left(-\frac{N}{6}\phi'(t_n) + \frac{M^2}{4}\right) / \phi'(t_n)\right] ||\bar{X} - X_n||^3 \\ &\le \left[-\left(-\frac{N}{6}\phi'(t_0) + \frac{M^2}{4}\right) / \phi'(t_n)\right] ||\bar{X} - X_n||^3 \\ &= \left[-\left(\frac{2N}{3B_0} + M^2\right) \frac{1}{4\phi'(t_n)}\right] ||\bar{X} - X_n||^3 \\ &\le -\frac{K^2}{4} \frac{1}{\phi'(t_n)} ||\bar{X} - X_n||^3. \end{aligned}$$

So,

$$\begin{aligned} ||\bar{X} - X_{n+1}|| &= ||\Theta_n \Gamma_n \tilde{F}_n|| \\ &\le \frac{\frac{K^2}{4} \frac{1}{(\phi'(t_n))^2}}{1 - \frac{K}{2} \frac{\phi(t_n)}{(\phi'(t_n))^2}} ||\bar{X} - X_n||^3. \end{aligned} \tag{2.49}$$

If $r_1 < r_2$, then $\bar{X} \in S(X_0, r_2)$. So $\exists \delta \in [0, 1)$ such that

$$||\bar{X} - X_0|| = \delta r_2 = \delta(r_2 - t_0). \tag{2.50}$$

If $r_1 = r_2$, then $\bar{X} \in \overline{S(X_0, r_1)}$. So $\exists \delta \in [0, 1]$ such that (2.50) holds. We claim that

$$||\bar{X} - X_k|| \le \delta^{3^k}(r_2 - t_k), \quad k = 0, 1, 2, \ldots. \tag{2.51}$$

In fact, this is true for $k = 0$ by (2.50). Suppose it is true for all $k \leq n$, then by (2.48) and (2.49), we have

$$\begin{aligned} ||\bar{X} - X_{n+1}|| &\leq \frac{\frac{K^2}{4}\frac{1}{(\phi'(t_n))^2}}{1 - \frac{K}{2}\frac{\phi(t_n)}{(\phi'(t_n))^2}}\delta^{3^{n+1}}(r_2 - t_n)^3 \\ &= \delta^{3^{n+1}}(r_2 - t_{n+1}). \end{aligned}$$

So, (2.51) is true. Now let $k \to \infty$. If $r_1 < r_2$, then $\delta < 1$ implies $||\bar{X} - X^*|| = 0$. When $r_1 = r_2$, then $t_k \to r_2$ and $\delta \leq 1$ imply $||\bar{X} - X^*|| = 0$. Thus, $\bar{X} = X^*$. □

Chapter 3

The Durand-Kerner Method and Aberth Method

3.1 Introduction

Durand-Kerner's method and Aberth's method are two major concurrent iterative methods for finding all zeros of a polynomial simultaneously (see Durand [16], Dochev [12], Kerner [32], Yamamoto [57], Petković [41], Börsch-Supan [6], Ehrlich [17], Aberth [1], Gargantini & Henrici [22], Braess & Hadeler [7], Alefeld & Herzberger [2], Yamamoto, et al. [57] and Petković [41] for references). In this chapter, we first present a new derivation of Durand-Kerner's method by homotopy. The new derivation gives a geometric interpretation of Durand-Kerner's method. Then, a two-step iterative scheme is proposed which is equivalent to Durand-Kerner's method but no evaluation of the polynomial is necessary for each iteration after the first step. For Aberth's method, an equivalent form is proposed which requires no evaluation of the first derivative for each iteration. We then present two r-step Aberth's methods, both of them having $2r + 1$ as their rate of convergence.

3.2 New derivation of Durand-Kerner's method by homotopy

Let $P(x)$ be a monic polynomial of degree n with n distinct zeros (real or complex) $x_1^*, x_2^*, \ldots, x_n^*$ or

$$P(x) = \prod_{j=1}^{n}(x - x_j^*). \tag{3.1}$$

Durand [16] and Kerner [32] have independently proposed the following concurrent iterative method with quadratically convergent rate for finding all zeros of the polynomial in (3.1) simultaneously:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod_{j=1,j\neq i}^{n}(x_i^{(k)} - x_j^{(k)})}, \quad i = 1, 2, \ldots, n; \quad k = 0, 1, 2, \ldots. \tag{3.2}$$

Durand-Kerner's method always converges in practice, see Fraigniaud [18], although there is no rigorous proof of it for $n \geq 3$. Many authors conjecture that Durand-Kerner's method converges for almost all starting points $\{x_i^{(0)}\}_{i=1}^{n}$ in C^n. One of the main interests of Durand-Kerner's method lies in its inherent parallelism.

For $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$, let

$$P_0(x) = \prod_{j=1}^{n}(x - x_j^{(0)}) \tag{3.3}$$

and write

$$H(x,t) = (1-t)P_0(x) + tP(x). \tag{3.4}$$

If there are n smooth curves $x_i(t)$, $i = 1, 2, \ldots, n$, which connect the initial approximations $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ to the zeros $x_1^*, x_2^*, \ldots, x_n^*$ of $P(x)$ respectively, then

$$H(x_i(t), t) \equiv (1-t)P_0(x_i(t)) + tP(x_i(t)) = 0, \quad i = 1, 2, \ldots, n. \tag{3.5}$$

Differentiating $H(x_i(t), t)$ yields

$$H_x \cdot x_i'(t) + H_t = 0, \quad i = 1, 2 \ldots, n, \tag{3.6}$$

where H_x and H_t are the partial derivatives of $H(x,t)$ with respect to x and t respectively. From (3.6), we have the following initial value problem

$$\begin{cases} x_i'(t) &= \frac{-H_t(x_i(t),t)}{H_x(x_i(t),t)}, \\ x_i(0) &= x_i^{(0)} \end{cases}$$

for $i = 1, 2, \ldots, n$. By (3.5), we have

$$\begin{cases} x_i'(t) &= \frac{P_0(x_i(t))-P(x_i(t))}{(1-t)P_0'(x_i(t))+tP'(x_i(t))}, \\ x_i(0) &= x_i^{(0)} \end{cases}$$

for $i = 1, 2, \ldots, n$ and $x_i(1) = x_i^*$, $i = 1, 2, \ldots, n$.

Now we use Euler's one-step formula to approximate $x_i(1)$, that is,

$$x_i(1) \approx x_i(0) + x_i'(0)$$

where

$$x_i'(0) = \frac{P_0(x_i^{(0)}) - P(x_i^{(0)})}{P_0'(x_i^{(0)})} = -\frac{P(x_i^{(0)})}{P_0'(x_i^{(0)})}.$$

Denote

$$x_i^{(1)} = x_i(0) + x_i'(0),$$

then

$$x_i^{(1)} = x_i^{(0)} - \frac{P(x_i^{(0)})}{P_0'(x_i^{(0)})}, \quad i = 1, 2, \ldots, n,$$

where $P_0'(x_i^{(0)}) = \prod_{j=1, j\neq i}^{n}(x_i^{(0)} - x_j^{(0)})$. In general, we have

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{P_k'(x_i^{(k)})}, \quad i = 1, 2, \ldots, n; \quad k = 0, 1, 2, \ldots, \tag{3.7}$$

where $P_k'(x_i^{(k)}) = \prod_{j=1, j\neq i}^{n}(x_i^{(k)} - x_j^{(k)})$. This is exactly the Durand-Kerner method in (3.2).

3.3 Two-step iterative scheme of Durand-Kerner's method

To use the Durand-Kerner method in (3.2), we must evaluate $P(x_i^{(k)})$ at each iterative step. In this section we propose the following two-step iterative scheme:

$$x_i^{(k+2)} = x_i^{(k+1)} - (x_i^{(k+1)} - x_i^{(k)})\left[\sum_{j=1, j\neq i}^{n} \frac{x_j^{(k)} - x_j^{(k+1)}}{x_i^{(k+1)} - x_j^{(k)}}\right]\left[\prod_{j=1, j\neq i}^{n} \frac{x_i^{(k+1)} - x_j^{(k)}}{x_i^{(k+1)} - x_j^{(k+1)}}\right] \tag{3.8}$$

for $i = 1, 2, \ldots, n$; $k = 0, 1, 2, \ldots$, where $x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}$ are initial approximations and

$$x_i^{(1)} = x_i^{(0)} - \frac{P(x_i^{(0)})}{\prod_{j=1, j\neq i}^{n}(x_i^{(0)} - x_j^{(0)})}, \quad i = 1, 2, \ldots, n.$$

We will show, in the following, that this iterative scheme is equivalent to the Durand-Kerner method in (3.2). With this scheme, the evaluation of $P(x_i^{(k)})$ for each iteration becomes unnecessary after the first step in (3.8). This iterative scheme is very useful when it is applied to solve the eigenvalue problem by the homotopy method where the evaluation of the polynomial is quite time-consuming.

Theorem 3.3.1 *The two-step iterative scheme in (3.8) is equivalent to the Durand-Kerner method in (3.2).*

Proof By Lagrange's interpolation formula over the nodes

$$(x_1^{(k)}, P(x_1^{(k)})), (x_2^{(k)}, P(x_2^{(k)})), \ldots, (x_n^{(k)}, P(x_n^{(k)}))$$

and Durand-Kerner's iteration in (3.2), we have

$$\begin{aligned} P(x) &= \sum_{l=1}^{n} \left[P(x_l^{(k)}) \prod_{j=1, j\neq l}^{n} \frac{x - x_j^{(k)}}{x_l^{(k)} - x_j^{(k)}} \right] + \prod_{j=1}^{n}(x - x_j^{(k)}) \\ &= \sum_{l=1}^{n} \left[(x_l^{(k)} - x_l^{(k+1)}) \prod_{j=1, j\neq l}^{n} (x - x_j^{(k)}) \right] + \prod_{j=1}^{n}(x - x_j^{(k)}). \end{aligned}$$

With $x = x_i^{(k+1)}$, the above equation becomes

$$\begin{aligned} P(x_i^{(k+1)}) &= \sum_{l=1}^{n} \left[(x_l^{(k)} - x_l^{(k+1)}) \prod_{j=1, j\neq l}^{n} (x_i^{(k+1)} - x_j^{(k)}) \right] + \prod_{j=1}^{n}(x_i^{(k+1)} - x_j^{(k)}) \\ &= \sum_{l=1, l\neq i}^{n} \left[(x_l^{(k)} - x_l^{(k+1)}) \prod_{j=1, j\neq l}^{n} (x_i^{(k+1)} - x_j^{(k)}) \right] \\ &= \left[\sum_{l=1, l\neq i}^{n} \frac{x_l^{(k)} - x_l^{(k+1)}}{x_i^{(k+1)} - x_l^{(k)}} \right] \prod_{j=1}^{n}(x_i^{(k+1)} - x_j^{(k)}). \end{aligned}$$

It follows that

$$\begin{aligned} x_i^{(k+2)} &= x_i^{(k+1)} - \frac{P(x_i^{(k+1)})}{\prod_{j=1, j\neq i}^{n}(x_i^{(k+1)} - x_j^{(k+1)})} \\ &= x_i^{(k+1)} - (x_i^{(k+1)} - x_i^{(k)}) \left[\sum_{l=1, l\neq i}^{n} \frac{x_l^{(k)} - x_l^{(k+1)}}{x_i^{(k+1)} - x_l^{(k)}} \right] \left[\prod_{j=1, j\neq i}^{n} \frac{x_i^{(k+1)} - x_j^{(k)}}{x_i^{(k+1)} - x_j^{(k+1)}} \right]. \end{aligned}$$

□

3.4 An equivalent form of Aberth's method

In [1], Aberth proposed the following concurrent iterative method with cubically convergent rate for finding all zeros of a monic polynomial in (3.1) simultaneously:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{1}{\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{j=1,j\neq i}^{n} \frac{1}{x_i^{(k)} - x_j^{(k)}}} \tag{3.9}$$

for $i = 1, 2, \ldots, n$; $k = 0, 1, 2, \ldots$. Actually, in [17], Ehrlich proposed an iteration similar to the one in (3.9). Like Durand-Kerner's method, Aberth's method always converges in practice although no rigorous proof of this property exists.

To use Aberth's method in (3.9), the evaluation of the first derivative of the polynomial for each iterative step is inevitable. In this section we will show that the Aberth method in (3.9) can actually be rewritten in the following form:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{d_i^{(k)}}{1 - \sum_{j=1,j\neq i}^{n} \frac{d_j^{(k)}}{x_j^{(k)} - x_i^{(k)}}}, \quad i = 1, 2, \ldots, n; \quad k = 0, 1, 2, \ldots, \tag{3.10}$$

where

$$d_i^{(k)} = \frac{P(x_i^{(k)})}{\prod_{j=1,j\neq i}^{n} (x_i^{(k)} - x_j^{(k)})}, \quad i = 1, 2, \ldots, n; \quad k = 0, 1, 2, \ldots. \tag{3.11}$$

In this form, the evaluation of the first derivative of the polynomial for each iterative step can be avoided.

Theorem 3.4.1 *The Aberth method in (3.9) is equivalent to the iterative formula in (3.10).*

Proof By (3.11) and Lagrange's interpolation formula over the nodes

$$(x_1^{(k)}, P(x_1^{(k)})), (x_2^{(k)}, P(x_2^{(k)})), \ldots, (x_n^{(k)}, P(x_n^{(k)})),$$

we have

$$\begin{aligned} P(x) &= \sum_{l=1}^{n} \left[P(x_l^{(k)}) \prod_{j=1,j\neq l}^{n} \frac{x - x_j^{(k)}}{x_l^{(k)} - x_j^{(k)}} \right] + \prod_{j=1}^{n} (x - x_j^{(k)}) \\ &= \sum_{l=1}^{n} \left[d_l^{(k)} \prod_{j=1,j\neq l}^{n} (x - x_j^{(k)}) \right] + \prod_{j=1}^{n} (x - x_j^{(k)}) \end{aligned} \tag{3.12}$$

and

$$P'(x) = \sum_{l=1}^{n} \left[d_l^{(k)} \sum_{r=1, r\neq l}^{n} \prod_{j=1, j\neq l, r}^{n} (x - x_j^{(k)}) \right] + \sum_{l=1}^{n} \prod_{j=1, j\neq l}^{n} (x - x_j^{(k)}). \tag{3.13}$$

Letting $x = x_i^{(k)}$ in (3.12) and (3.13), we have

$$P(x_i^{(k)}) = d_i^{(k)} \prod_{j=1, j\neq i}^{n} (x_i^{(k)} - x_j^{(k)})$$

and

$$\begin{aligned} P'(x_i^{(k)}) &= \sum_{l=1}^{n} \left[d_l^{(k)} \sum_{r=1, r\neq l}^{n} \prod_{j=1, j\neq l, r}^{n} (x_i^{(k)} - x_j^{(k)}) \right] + \prod_{j=1, j\neq i}^{n} (x_i^{(k)} - x_j^{(k)}) \\ &= d_i^{(k)} \sum_{r=1, r\neq i}^{n} \prod_{j=1, j\neq i, r}^{n} (x_i^{(k)} - x_j^{(k)}) + \sum_{l=1, l\neq i}^{n} d_l^{(k)} \prod_{j=1, j\neq l, i}^{n} (x_i^{(k)} - x_j^{(k)}) \\ &\quad + \prod_{j=1, j\neq i}^{n} (x_i^{(k)} - x_j^{(k)}) \\ &= d_i^{(k)} \sum_{l=1, l\neq i}^{n} \prod_{j=1, j\neq i, l}^{n} (x_i^{(k)} - x_j^{(k)}) + \sum_{l=1, l\neq i}^{n} d_l^{(k)} \prod_{j=1, j\neq l, i}^{n} (x_i^{(k)} - x_j^{(k)}) \\ &\quad + \prod_{j=1, j\neq i}^{n} (x_i^{(k)} - x_j^{(k)}). \end{aligned}$$

So

$$\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} = \sum_{l=1, l\neq i}^{n} \frac{1}{x_i^{(k)} - x_l^{(k)}} + \frac{1}{d_i^{(k)}} \sum_{l=1, l\neq i}^{n} \frac{d_l^{(k)}}{x_i^{(k)} - x_l^{(k)}} + \frac{1}{d_i^{(k)}},$$

or

$$\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{l=1, l\neq i}^{n} \frac{1}{x_i^{(k)} - x_l^{(k)}} = \frac{1}{d_i^{(k)}} \left[1 - \sum_{l=1, l\neq i}^{n} \frac{d_l^{(k)}}{x_l^{(k)} - x_i^{(k)}} \right].$$

Thus,

$$\frac{1}{\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{l=1, l\neq i}^{n} \frac{1}{x_i^{(k)} - x_l^{(k)}}} = \frac{d_i^{(k)}}{1 - \sum_{l=1, l\neq i}^{n} \frac{d_l^{(k)}}{x_l^{(k)} - x_i^{(k)}}}$$

and we have proved the assertion of the theorem. □

3.5 *R*-step Aberth's methods

In this section we present derivations of Aberth's method and its equivalent form in (3.10). Those derivations provide motivations for considering the corresponding r-step Aberth's methods with $(2r+1)$-order convergence rate.

For the monic polynomial $P(x)$ in (3.1), we have

$$\frac{P'(x)}{P(x)} = \sum_{j=1}^{n} \frac{1}{x - x_j^*}.$$

Replacing x by $x_i^{(k)}$ yields

$$\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} = \sum_{j=1, j\neq i}^{n} \frac{1}{x_i^{(k)} - x_j^*} + \frac{1}{x_i^{(k)} - x_i^*}.$$

It follows that

$$x_i^* = x_i^{(k)} - \frac{1}{\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{j=1, j\neq i}^{n} \frac{1}{x_i^{(k)} - x_j^*}} \tag{3.14}$$

for $i = 1, 2, \ldots, n$. Let $x_j^* = x_j^{(k)}$ for $j \neq i$ and denote the right hand side by $x_i^{(k+1)}$, we obtain the Aberth method given in (3.9).

The formulation of the equation in (3.14) may be considered as a fixed-point problem which suggests the consideration of the following r-step Aberth's method:

$$\begin{cases} y_i^{(k,0)} &= x_i^{(k)}, \\ y_i^{(k,l+1)} &= x_i^{(k)} - \dfrac{1}{\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{j=1, j\neq i}^{n} \frac{1}{x_i^{(k)} - y_j^{(k,l)}}}, \quad l = 0, 1, \ldots, r-1, \\ x_i^{(k+1)} &= y_i^{(k,r)} \end{cases} \tag{3.15}$$

for $i = 1, 2, \ldots, n;\quad k = 0, 1, 2, \ldots$.

It can easily be shown by mathematical induction that the order of convergence of the r-step Aberth's method in (3.15) is $2r + 1$ by deriving

$$y_i^{(k,l+1)} - x_i^* = O\left((x_i^{(k)} - x_i^*)^2 \sum_{j=1, j\neq i}^{n} (y_j^{(k,l)} - x_j^*)\right)$$

for $i = 1, 2, \ldots, n;\quad l = 0, 1, \ldots, r-1$.

Comparing to Aberth's method in (3.9), the r-step Aberth's method in (3.15) requires no extra work on evaluating $P'(x_i^{(k)})$ and $P(x_i^{(k)})$ while the same formula in (3.15) is reused r times with the same $P'(x_i^{(k)})$ and $P(x_i^{(k)})$.

To derive the equivalent form in (3.10) of the Aberth method, recall that,

$$P(x) = \left[1 - \sum_{l=1}^{n} \frac{d_l^{(k)}}{x_l^{(k)} - x}\right] \prod_{j=1}^{n} (x - x_j^{(k)})$$

in (3.12). Let $x = x_i^*$, $i = 1, 2, \ldots, n$. With the assumption $x_i^{(k)} \neq x_i^*$, $i = 1, 2, \ldots, n$, we have

$$1 - \sum_{j=1}^{n} \frac{d_j^{(k)}}{x_j^{(k)} - x_i^*} = 0,$$

or

$$\frac{d_i^{(k)}}{x_i^{(k)} - x_i^*} = 1 - \sum_{j=1, j \neq i}^{n} \frac{d_j^{(k)}}{x_j^{(k)} - x_i^*}.$$

Thus,

$$x_i^* = x_i^{(k)} - \frac{d_i^{(k)}}{1 - \sum_{j=1, j \neq i}^{n} \frac{d_j^{(k)}}{x_j^{(k)} - x_i^*}}, \quad i = 1, 2, \ldots, n. \tag{3.16}$$

Let $x_i^* = x_i^{(k)}$ and denote the right hand side of (3.16) by $x_i^{(k+1)}$, we achieve the equivalent form of the Aberth method in (3.10).

One may also consider the formulation in the equation in (3.16) as a fixed-point problem. This suggests the consideration of the following r-step Aberth's method:

$$\begin{cases} y_i^{(k,0)} &= x_i^{(k)}, \\ y_i^{(k,l+1)} &= x_i^{(k)} - \dfrac{d_i^{(k)}}{1 - \sum_{j=1, j \neq i}^{n} \frac{d_j^{(k)}}{x_j^{(k)} - y_i^{(k,l)}}}, \quad l = 0, 1, \ldots, r-1, \\ x_i^{(k+1)} &= y_i^{(k,r)} \end{cases} \tag{3.17}$$

for $i = 1, 2, \ldots, n$; $k = 0, 1, 2, \ldots$, where $d_i^{(k)}$ is the same as in (3.11).

Similarly, it can easily be shown that the order of convergence of the r-step Aberth's method in (3.17) is $2r + 1$. Compared with the equivalent form of the Aberth method in (3.10), the r-step Aberth method in (3.17) requires no extra work in evaluating $d_i^{(k)}$ but reuses the same formula in (3.17) r times with the same $d_i^{(k)}$.

Chapter 4

A New Method with a Quadratic Convergence Rate for Simultaneously Finding Polynomial Zeros

4.1 Introduction

Durand-Kerner's method is an iterative algorithm for finding all zeros of a monic polynomial simultaneously with quadratic convergence rate. It was first used by Weierstrass [55] and was later proposed independently by Durand [16], Dochev [12] and Kerner [32] (see also Yamamoto [57] and Petković [41]). There are many algorithms of Durand-Kerner-type with higher order convergence rate, which can be found in Börsch-Supan [6], Ehrlich [17], Aberth [1], Gargantini & Henrici [22], Braess & Hadeler [7], Alefeld & Herzberger [2], Yamamoto, et al. [57] and Petković [41]. However, none of the methods mentioned above converges monotonically for finding real zeros of polynomials. In this chapter we propose new iterative methods which converge monotonically for both simple and multiple real zeros of polynomials. Our new iterative methods converge quadratically, and most importantly, they require no evaluation of derivatives.

4.2 A new method with quadratic convergence rate and monotonic convergence

Consider a monic polynomial of degree $n \geq 3$,

$$P(z) = \prod_{i=1}^{n}(z - \lambda_i), \tag{4.1}$$

with simple real zeros

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n. \tag{4.2}$$

First, we construct a pair of sequences $\{x_i^{(k)}\}_{i=1}^n$ and $\{y_i^{(k)}\}_{i=1}^n$ by the following iterations with a pair of initial values $\{x_i^{(0)}\}_{i=1}^n$ and $\{y_i^{(0)}\}_{i=1}^n$:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n}(x_i^{(k)} - y_j^{(k)})}, \quad i = 1, \ldots, n, \tag{4.3}$$

$$y_i^{(k+1)} = y_i^{(k)} - \frac{P(y_i^{(k)})}{\prod_{j=1}^{i-1}(y_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n}(y_i^{(k)} - y_j^{(k)})}, \quad i = 1, \ldots, n. \tag{4.4}$$

for $k = 0, 1, 2, \ldots$. These two sequences can be used to approximate all zeros of the polynomial $P(z)$ simultaneously. In the following theorem, we show that the iterations defined in (4.3) and (4.4) are monotonically convergent.

Theorem 4.2.1 *For the polynomial $P(z)$ in (4.1), if the initial values $\{x_i^{(0)}\}_{i=1}^n$ and $\{y_i^{(0)}\}_{i=1}^n$ satisfy*

$$x_1^{(0)} \leq \lambda_1 \leq y_1^{(0)} < x_2^{(0)} \leq \lambda_2 \leq y_2^{(0)} < \cdots < x_n^{(0)} \leq \lambda_n \leq y_n^{(0)}, \tag{4.5}$$

then

(a) *The sequence $\{x_i^{(k)}\}_{k=0}^{\infty}$ generated by (4.3) converges to λ_i monotonically. That is,*

$$x_i^{(0)} < x_i^{(1)} < \cdots < x_i^{(k)} \xrightarrow{k\to\infty} \lambda_i, \quad i = 1, \ldots, n. \tag{4.6}$$

(b) *The sequence* $\{y_i^{(k)}\}_{k=0}^{\infty}$ *generated by* (4.4) *converges to* λ_i *monotonically. That is,*

$$\lambda_i \overset{k\to\infty}{\longleftarrow} y_i^{(k)} < y_i^{(k-1)} < \cdots < y_i^{(1)} < y_i^{(0)}, \quad i = 1, \ldots, n. \tag{4.7}$$

Proof Let $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i = 1, \ldots, n$. We have, by (4.3),

$$\begin{aligned} \varepsilon_i^{(k+1)} &= \varepsilon_i^{(k)} + (x_i^{(k)} - \lambda_i) \prod_{j=1}^{i-1} \frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - x_j^{(k)}} \prod_{j=i+1}^{n} \frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - y_j^{(k)}} \\ &= \varepsilon_i^{(k)} \left[1 - \prod_{j=1}^{i-1} \frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - x_j^{(k)}} \prod_{j=i+1}^{n} \frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - y_j^{(k)}} \right]. \end{aligned} \tag{4.8}$$

For $k = 0$, by (4.5),

$$0 < \varepsilon_i^{(1)} = \varepsilon_i^{(0)} \left[1 - \prod_{j=1}^{i-1} \frac{x_i^{(0)} - \lambda_j}{x_i^{(0)} - x_j^{(0)}} \prod_{j=i+1}^{n} \frac{x_i^{(0)} - \lambda_j}{x_i^{(0)} - y_j^{(0)}} \right] < \varepsilon_i^{(0)}$$

when $\varepsilon_i^{(0)} \neq 0$, and $\varepsilon_i^{(1)} = 0$ for $\varepsilon_i^{(0)} = 0$. So

$$x_i^{(0)} < x_i^{(1)} < \lambda_i \quad \text{or} \quad x_i^{(0)} = x_i^{(1)} = \lambda_i.$$

Similarly, by (4.4), we have

$$\begin{aligned} \delta_i^{(k+1)} &= \delta_i^{(k)} - (y_i^{(k)} - \lambda_i) \prod_{j=1}^{i-1} \frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - x_j^{(k)}} \prod_{j=i+1}^{n} \frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - y_j^{(k)}} \\ &= \delta_i^{(k)} \left[1 - \prod_{j=1}^{i-1} \frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - x_j^{(k)}} \prod_{j=i+1}^{n} \frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - y_j^{(k)}} \right]. \end{aligned} \tag{4.9}$$

By (4.5),

$$0 < \delta_i^{(1)} = \delta_i^{(0)} \left[1 - \prod_{j=1}^{i-1} \frac{y_i^{(0)} - \lambda_j}{y_i^{(0)} - x_j^{(0)}} \prod_{j=i+1}^{n} \frac{y_i^{(0)} - \lambda_j}{y_i^{(0)} - y_j^{(0)}} \right] < \delta_i^{(0)}$$

when $\delta_i^{(0)} \neq 0$, and $\delta_i^{(1)} = 0$ for $\delta_i^{(0)} = 0$. So,

$$\lambda_i < y_i^{(1)} < y_i^{(0)} \quad \text{or} \quad y_i^{(1)} = y_i^{(0)} = \lambda_i.$$

Therefore, by mathematical induction, (4.6) and (4.7) can easily be obtained from (4.8) and (4.9). □

Theorem 4.2.2 *The convergence rate of the iterations in (4.3) and (4.4) is quadratic for simple zeros of the polynomial* $P(z)$ *in (4.1).*

Proof Let $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i = 1, \ldots, n$. Then, by (4.8), we have

$$\varepsilon_i^{(k+1)} = \varepsilon_i^{(k)} \frac{N_1}{D_1}$$

where

$$\begin{aligned}
N_1 &= \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n} (x_i^{(k)} - y_j^{(k)}) - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j) \prod_{j=i+1}^{n} (x_i^{(k)} - \lambda_j) \\
&= \prod_{j=1}^{i-1}[(x_i^{(k)} - \lambda_j) + \varepsilon_j^{(k)}] \prod_{j=i+1}^{n} [(x_i^{(k)} - \lambda_j) - \delta_j^{(k)}] \\
&\quad - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j) \prod_{j=i+1}^{n} (x_i^{(k)} - \lambda_j) \\
&= \sum_{l=1}^{i-1}[\varepsilon_l^{(k)} \prod_{j\neq i,l}^{n} (x_i^{(k)} - \lambda_j)] - \sum_{l=i+1}^{n} [\delta_l^{(k)} \prod_{j\neq i,l}^{n} (x_i^{(k)} - \lambda_j)] \\
&\quad + O(\varepsilon_i^{(k)}\delta_j^{(k)} + \varepsilon_i^{(k)}\varepsilon_j^{(k)} + \delta_i^{(k)}\delta_j^{(k)})
\end{aligned}$$

and

$$D_1 = \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n} (x_i^{(k)} - y_j^{(k)}).$$

Since $\lambda_1, \ldots, \lambda_n$ are distinct, $\forall k$, $\exists d$ and D such that

$$d < |x_i^{(k)} - x_j^{(k)}|, |x_i^{(k)} - y_j^{(k)}|, |y_i^{(k)} - y_j^{(k)}| < D, \quad \forall i \neq j, \ \forall k;$$

$$d < |x_i^{(k)} - \lambda_j|, |y_i^{(k)} - \lambda_j| < D, \quad \forall i \neq j, \ \forall k.$$

Let $\varepsilon^{(k)} = \max_{i=1,\ldots,n}\{\varepsilon_i^{(k)}, \delta_i^{(k)}\}$, then

$$\begin{aligned}
\varepsilon_i^{(k+1)} &\leq \varepsilon_i^{(k)} \left|\frac{N_1}{D_1}\right| \\
&\leq (\varepsilon^{(k)})^2 (n-1) \frac{D^{n-2}}{d^{n-1}} + O\big((\varepsilon^{(k)})^3\big).
\end{aligned}$$

Similarly, we can obtain

$$\delta_i^{(k+1)} \leq (\varepsilon^{(k)})^2 (n-1) \frac{D^{n-2}}{d^{n-1}} + O\big((\varepsilon^{(k)})^3\big).$$

Thus,

$$\varepsilon^{(k+1)} \leq (n-1) \frac{D^{n-2}}{d^{n-1}} (\varepsilon^{(k)})^2 + O\big((\varepsilon^{(k)})^3\big).$$

□

In the rest of this section, we introduce an alternative scheme for the iterations in (4.3) and (4.4) which reduces the computational cost of each iteration.

For $i = 1, \ldots, n$, consider the midpoint $c_i^{(k)} = (x_i^{(k)} + y_i^{(k)})/2$ of the interval $[x_i^{(k)}, y_i^{(k)}]$. For a zero λ_i of $P(z)$, we may detect if $\lambda_i \in [x_i^{(k)}, c_i^{(k)}]$ or $\lambda_i \in [c_i^{(k)}, y_i^{(k)}]$. If $\lambda_i \in [x_i^{(k)}, c_i^{(k)}]$, we use (4.4), othervise we apply (4.3). That is:

(1) If $c_i^{(k)} > \lambda_i$, then

$$x_i^{(k+1)} = x_i^{(k)}; \tag{4.10}$$

$$y_i^{(k+1)} = c_i^{(k)} - \frac{P(c_i^{(k)})}{\prod_{j=1}^{i-1}(c_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n}(c_i^{(k)} - y_j^{(k)})}. \tag{4.11}$$

(2) If $c_i^{(k)} < \lambda_i$, then

$$x_i^{(k+1)} = c_i^{(k)} - \frac{P(c_i^{(k)})}{\prod_{j=1}^{i-1}(c_i^{(k)} - x_j^{(k)}) \prod_{j=i+1}^{n}(c_i^{(k)} - y_j^{(k)})}; \tag{4.12}$$

$$y_i^{(k+1)} = y_i^{(k)}. \tag{4.13}$$

It is easy to see that the above scheme converges monotonically for approximating the zeros of $P(z)$.

4.3 Multiple zeros and clusters of zeros

Theorem 4.2.2 ensures the iterations in (4.3) and (4.4) converge quadratically only for simple zeros. In this section we will consider the multiple-zero case. For a monic polynomial of degree $n \geq 3$,

$$P(z) = \prod_{i=1}^{m}(z - \lambda_i)^{n_i} \tag{4.14}$$

with zeros

$$\lambda_1 < \lambda_2 < \cdots < \lambda_m \quad \text{and} \quad \sum_{i=1}^{m} n_i = n, \tag{4.15}$$

we can modify the iterations in (4.3) and (4.4) as follows:

$$x_i^{(k+1)} = x_i^{(k)} + \sqrt[n_i]{\left|\frac{P(x_i^{(k)})}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k)})^{n_j} \prod_{j=i+1}^{m}(x_i^{(k)} - y_j^{(k)})^{n_j}}\right|} \equiv A_{n_i}(x_i^{(k)})\ ; \tag{4.16}$$

$$y_i^{(k+1)} = y_i^{(k)} - \sqrt[n_i]{\frac{P(y_i^{(k)})}{\prod_{j=1}^{i-1}(y_i^{(k)} - x_j^{(k)})^{n_j} \prod_{j=i+1}^{m}(y_i^{(k)} - y_j^{(k)})^{n_j}}} \equiv B_{n_i}(y_i^{(k)}) \,, \tag{4.17}$$

for $i = 1, \ldots, m$ and $k = 0, 1, 2, \ldots$. These modified iterations converge monotonically for the multiple zeros of $P(z)$ as shown in the following theorem.

Theorem 4.3.1 *If the initial values* $\{x_i^{(0)}\}_{i=1}^m$ *and* $\{y_i^{(0)}\}_{i=1}^m$ *satisfy :*

$$x_1^{(0)} \le \lambda_1 \le y_1^{(0)} < x_2^{(0)} \le \lambda_2 \le y_2^{(0)} < \cdots < x_m^{(0)} \le \lambda_m \le y_m^{(0)}, \tag{4.18}$$

then

(a) *The sequence* $\{x_i^{(k)}\}_{k=0}^\infty$ *generated by* (4.16) *converges to* λ_i *monotonically. That is,*

$$x_i^{(0)} < x_i^{(1)} < \cdots < x_i^{(k)} \overset{k\to\infty}{\longrightarrow} \lambda_i, \quad i = 1, \ldots, m. \tag{4.19}$$

(b) *The sequence* $\{y_i^{(k)}\}_{k=0}^\infty$ *generated by* (4.17) *converges to* λ_i *monotonically. That is,*

$$\lambda_i \overset{k\to\infty}{\longleftarrow} y_i^{(k)} < y_i^{(k-1)} < \cdots < y_i^{(1)} < y_i^{(0)}, \quad i = 1, \ldots, m. \tag{4.20}$$

Proof Let $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i = 1, \ldots, m$. It follows from (4.16),

$$\begin{aligned} \varepsilon_i^{(k+1)} &= \varepsilon_i^{(k)} - |x_i^{(k)} - \lambda_i| \sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - x_j^{(k)}}\right)^{n_j} \prod_{j=i+1}^{m}\left(\frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - y_j^{(k)}}\right)^{n_j}} \\ &= \varepsilon_i^{(k)}\left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - x_j^{(k)}}\right)^{n_j} \prod_{j=i+1}^{m}\left(\frac{x_i^{(k)} - \lambda_j}{x_i^{(k)} - y_j^{(k)}}\right)^{n_j}}\right]. \end{aligned} \tag{4.21}$$

For $k = 0$, by (4.18),

$$0 < \varepsilon_i^{(1)} = \varepsilon_i^{(0)}\left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{x_i^{(0)} - \lambda_j}{x_i^{(0)} - x_j^{(0)}}\right)^{n_j} \prod_{j=i+1}^{m}\left(\frac{x_i^{(0)} - \lambda_j}{x_i^{(0)} - y_j^{(0)}}\right)^{n_j}}\right] < \varepsilon_i^{(0)}$$

when $\delta_i^{(0)} \neq 0$, and $\varepsilon_i^{(1)} = 0$ for $\varepsilon_i^{(0)} = 0$. So,

$$x_i^{(0)} < x_i^{(1)} < \lambda_i \;\text{ or }\; x_i^{(0)} = x_i^{(1)} = \lambda_i.$$

Similarly, by (4.17), we have

$$\begin{aligned}\delta_i^{(k+1)} &= \delta_i^{(k)} - (y_i^{(k)} - \lambda_i)\sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - x_j^{(k)}}\right)^{n_j}\prod_{j=i+1}^{m}\left(\frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - y_j^{(k)}}\right)^{n_j}} \\ &= \delta_i^{(k)}\left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - x_j^{(k)}}\right)^{n_j}\prod_{j=i+1}^{m}\left(\frac{y_i^{(k)} - \lambda_j}{y_i^{(k)} - y_j^{(k)}}\right)^{n_j}}\right] \qquad (4.22)\end{aligned}$$

and by (4.18),

$$0 < \delta_i^{(1)} = \delta_i^{(0)}\left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1}\left(\frac{y_i^{(0)} - \lambda_j}{y_i^{(0)} - x_j^{(0)}}\right)^{n_j}\prod_{j=i+1}^{m}\left(\frac{y_i^{(0)} - \lambda_j}{y_i^{(0)} - y_j^{(0)}}\right)^{n_j}}\right] < \delta_i^{(0)}$$

if $\delta_i^{(0)} \neq 0$, and $\delta_i^{(1)} = 0$ when $\delta_i^{(0)} = 0$. So,

$$\lambda_i < y_i^{(1)} < y_i^{(0)} \quad \text{or} \quad y_i^{(1)} = y_i^{(0)} = \lambda_i.$$

Therefore, by mathematical induction, (4.19) and (4.20) can easily be obtained from (4.21) and (4.22). □

In the next theorem we show that the modified iterations in (4.16) and (4.17) have non-overshoot properties for polynomials with clusters of zeros. Namely, they never skip over n_i zeros of a polynomial. The non-overshoot properties are important when they are applied to find the eigenvalues of a symmetric matrix (see [36]).

Theorem 4.3.2 *If $P(z) = \prod_{i=1}^{n}(z - \lambda_i)$ has a cluster of zeros $\{\lambda_i^{(1)}, \lambda_i^{(2)}, \ldots, \lambda_i^{(n_i)}\}_{i=1}^{m}$ with $\lambda_i^{(1)} \leq \lambda_i^{(2)} \leq \cdots \leq \lambda_i^{(n_i)} < \lambda_{i+1}^{(1)}$ and the initial values $\{x_i\}_{i=1}^{m}$ and $\{y_i\}_{i=1}^{m}$ satisfy*

$$\begin{aligned} x_1 &\leq \lambda_1^{(1)} \leq \cdots \leq \lambda_1^{(n_1)} \leq y_1 < x_2 \leq \lambda_2^{(1)} \leq \cdots \\ &\leq \lambda_2^{(n_2)} \leq y_2 < \cdots < x_m \leq \lambda_m^{(1)} \leq \cdots \leq \lambda_m^{(n_m)} \leq y_m \qquad (4.23)\end{aligned}$$

where $\sum_{i=1}^{m} n_i = n$, then

$$x_i < A_1(x_i) < A_2(x_i) < \cdots < A_{n_i}(x_i) < \lambda_i^{(n_i)}, \quad i = 1, \ldots, m \qquad (4.24)$$

and

$$\lambda_i^{(1)} < B_{n_i}(y_i) < B_{n_i-1}(y_i) < \cdots < B_1(y_i) < y_i, \quad i = 1, \ldots, m \qquad (4.25)$$

where $A_{n_i}(x_i)$ and $B_{n_i}(y_i)$ are the same as in the iterations (4.16) and (4.17).

Proof By (4.16), (4.17) and (4.23), we have

$$\begin{aligned}
\lambda_i^{(n_i)} - A_{n_i}(x_i) &> (\lambda_i^{(n_i)} - x_i) - \\
&\quad \sqrt[n_i]{|x_i - \lambda_i^{(n_i)}|^{n_i} \prod_{j=1}^{i-1} \left(\frac{x_i - \lambda_j^{(1)}}{x_i - x_j}\right)^{n_j} \prod_{j=i+1}^{m} \left(\frac{x_i - \lambda_j^{(n_j)}}{x_i - y_j}\right)^{n_j}} \\
&= (\lambda_i^{(n_i)} - x_i) \left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1} \left(\frac{x_i - \lambda_j^{(1)}}{x_i - x_j}\right)^{n_j} \prod_{j=i+1}^{m} \left(\frac{x_i - \lambda_j^{(n_j)}}{x_i - y_j}\right)^{n_j}}\right] \\
&> 0;
\end{aligned}$$

$$\begin{aligned}
B_{n_i}(y_i) - \lambda_i^{(1)} &> (y_i - \lambda_i^{(1)}) - \\
&\quad \sqrt[n_i]{(y_i - \lambda_i^{(1)})^{n_i} \prod_{j=1}^{i-1} \left(\frac{y_i - \lambda_j^{(1)}}{y_i - x_j}\right)^{n_j} \prod_{j=i+1}^{m} \left(\frac{y_i - \lambda_j^{(n_j)}}{y_i - y_j}\right)^{n_j}} \\
&= (y_i - \lambda_i^{(1)}) \left[1 - \sqrt[n_i]{\prod_{j=1}^{i-1} \left(\frac{y_i - \lambda_j^{(1)}}{y_i - x_j}\right)^{n_j} \prod_{j=i+1}^{m} \left(\frac{y_i - \lambda_j^{(n_j)}}{y_i - y_j}\right)^{n_j}}\right] \\
&> 0.
\end{aligned}$$

and it is easy to check that

$$\frac{\partial A_{n_i}(x_i)}{\partial n_i} > 0 \quad \text{and} \quad \frac{\partial B_{n_i}(y_i)}{\partial n_i} < 0.$$

Therefore, (4.24) and (4.25) follow. □

Theorem 4.3.3 *The convergence rates of the modified iterations in (4.16) and (4.17) are both quadratic for an n_i-fold zero λ_i of $P(z)$.*

Proof It is similar to the proof of Theorem 4.2.2. □

4.4 Single-step methods

We call the iterations in (4.3) and (4.4) for simple zeros as well as the modified iterations in (4.16) and (4.17) for multiple zeros the *Total-step Methods.* In this section we consider their corresponding *Single-step Methods* which have the same convergence properties, but with faster convergence speed.

First consider the case of simple zeros. For polynomial $P(z)$, we define the following single-step iterations (denoted by *SSIMS*):

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)}) \prod_{j=i+1}^{n}(x_i^{(k)} - y_j^{(k)})}, \tag{4.26}$$

$$y_i^{(k+1)} = y_i^{(k)} - \frac{P(y_i^{(k)})}{\prod_{j=1}^{i-1}(y_i^{(k)} - x_j^{(k+1)}) \prod_{j=i+1}^{n}(y_i^{(k)} - y_j^{(k)})}, \tag{4.27}$$

where $k = 0, 1, 2, \ldots,$ and $i = 1, 2, \ldots, n$. As in Theorem 4.2.1, one can prove the following monotonic convergence theorem.

Theorem 4.4.1 *If the initial values* $\{x_i^{(0)}\}_{i=1}^n$ *and* $\{y_i^{(0)}\}_{i=1}^n$ *satisfy*

$$x_1^{(0)} \le \lambda_1 \le y_1^{(0)} < x_2^{(0)} \le \lambda_2 \le y_2^{(0)} < \cdots < x_n^{(0)} \le \lambda_n \le y_n^{(0)}, \tag{4.28}$$

then

(a) *The sequence* $\{x_i^{(k)}\}_{k=0}^{\infty}$ *generated by* (4.26) *converges to* λ_i *monotonically. That is,*

$$x_i^{(0)} < x_i^{(1)} < \cdots < x_i^{(k)} \stackrel{k\to\infty}{\longrightarrow} \lambda_i, \quad i = 1, \ldots, n. \tag{4.29}$$

(b) *The sequence* $\{y_i^{(k)}\}_{k=0}^{\infty}$ *generated by* (4.27) *converges to* λ_i *monotonically. That is,*

$$\lambda_i \stackrel{k\to\infty}{\longleftarrow} y_i^{(k)} < y_i^{(k-1)} < \cdots < y_i^{(1)} < y_i^{(0)}, \quad i = 1, \ldots, n. \tag{4.30}$$

For the convergence speed of the single-step iterations in (4.26) and (4.27), we will use the **R**-order of convergence concept given by Ortega and Rheinboldt (see [38]).

Theorem 4.4.2 *Let* $\sigma_n > 1$ *be the unique positive root of* $q_n(\sigma) = \sigma^n - \sigma - 1 = 0$. *Then for the* **R**-*order of (SSIMS), we have*

$$O_{\mathbf{R}}((SSIMS), \lambda_i) \ge 1 + \sigma_n.$$

Proof Let $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i = 1, \ldots, n$. Then, by (4.26), we have

$$\varepsilon_i^{(k+1)} = \varepsilon_i^{(k)} \frac{N_2}{D_2}$$

where

$$\begin{aligned} N_1 &= \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)}) \prod_{j=i+1}^{n} (x_i^{(k)} - y_j^{(k)}) - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j) \prod_{j=i+1}^{n} (x_i^{(k)} - \lambda_j) \\ &= \prod_{j=1}^{i-1}[(x_i^{(k)} - \lambda_j) + \varepsilon_j^{(k+1)}] \prod_{j=i+1}^{n} [(x_i^{(k)} - \lambda_j) - \delta_j^{(k)}] \\ &\quad - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j) \prod_{j=i+1}^{n} (x_i^{(k)} - \lambda_j) \\ &= \sum_{l=1}^{i-1}[\varepsilon_l^{(k+1)} \prod_{j\neq i,l}^{n} (x_i^{(k)} - \lambda_j)] - \sum_{l=i+1}^{n} [\delta_l^{(k)} \prod_{j\neq i,l}^{n} (x_i^{(k)} - \lambda_j)] \\ &\quad + O(\varepsilon_i^{(k+1)}\delta_j^{(k)} + \varepsilon_i^{(k+1)}\varepsilon_j^{(k+1)} + \delta_i^{(k)}\delta_j^{(k)}) \end{aligned}$$

and

$$D_1 = \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)}) \prod_{j=i+1}^{n} (x_i^{(k)} - y_j^{(k)}).$$

Since $\lambda_1, \ldots, \lambda_n$ are distinct, by (4.28), $\exists d$ and D such that

$$d < |x_i^{(k)} - x_j^{(k+1)}|, |x_i^{(k)} - y_j^{(k)}|, |y_i^{(k)} - y_j^{(k)}|, |y_i^{(k)} - x_j^{(k+1)}| < D, \quad \forall i \neq j, \quad \forall k;$$

$$d < |x_i^{(k)} - \lambda_j|, |y_i^{(k)} - \lambda_j| < D, \quad \forall i \neq j, \quad \forall k.$$

So,

$$\begin{aligned} |N_2| &\leq D^{n-2}\left[\sum_{j=1}^{i-1}\varepsilon_j^{(k+1)} + \sum_{j=i+1}^{n}\delta_j^{(k)}\right] + O(\varepsilon_i^{(k+1)}\delta_j^{(k)} + \varepsilon_i^{(k+1)}\varepsilon_j^{(k+1)} + \delta_i^{(k)}\delta_j^{(k)}); \\ |D_2| &\geq d^{n-1}. \end{aligned}$$

Thus, $\exists C_1 > 0$ such that

$$\begin{aligned} \varepsilon_i^{(k+1)} &\leq \varepsilon_i^{(k)}\left|\frac{N_2}{D_2}\right| \\ &\leq C_1\varepsilon_i^{(k)}\left[\sum_{j=1}^{i-1}\varepsilon_j^{(k+1)} + \sum_{j=i+1}^{n}\delta_j^{(k)}\right]. \end{aligned}$$

Similary, $\exists C_2 > 0$ such that

$$\delta_i^{(k+1)} \leq C_2\delta_i^{(k)}\left[\sum_{j=1}^{i-1}\varepsilon_j^{(k+1)} + \sum_{j=i+1}^{n}\delta_j^{(k)}\right].$$

Let

$$C = \max\{C_1, C_2\} \quad \text{and} \quad \beta = (n-1)C;$$

$$\eta_i^{(k)} = \beta\varepsilon_i^{(k)} \quad \text{and} \quad \rho_i^{(k)} = \beta\delta_i^{(k)}, \quad i = 1, \ldots, n.$$

Then,

$$\begin{aligned} \eta_i^{(k+1)} &\leq \frac{1}{n-1}\eta_i^{(k)} \left[\sum_{j=1}^{i-1} \eta_j^{(k+1)} + \sum_{j=i+1}^{n} \rho_j^{(k)}\right]; \\ \rho_i^{(k+1)} &\leq \frac{1}{n-1}\rho_i^{(k)} \left[\sum_{j=1}^{i-1} \eta_j^{(k+1)} + \sum_{j=i+1}^{n} \rho_j^{(k)}\right]. \end{aligned}$$

Since

$$\lim_{k\to\infty} \varepsilon_i^{(k)} = 0 \quad \text{and} \quad \lim_{k\to\infty} \delta_i^{(k)} = 0, \quad i = 1, \ldots, n,$$

we may assume there is a constant $\eta > 0$ such that

$$\eta_i^{(0)} \leq \eta < 1 \quad \text{and} \quad \rho_i^{(0)} \leq \eta < 1, \quad i = 1, \ldots, n.$$

By mathematical induction, we can obtain the following:

$$\begin{aligned} \eta_i^{(k+1)} &\leq \eta^{m_i^{(k+1)}}; \\ \rho_i^{(k+1)} &\leq \eta^{m_i^{(k+1)}} \end{aligned}$$

where

$$m^{(k)} = \begin{pmatrix} m_1^{(k)} \\ \vdots \\ m_n^{(k)} \end{pmatrix} \quad \text{with } m_i^{(0)} = 1, \quad i = 1, \ldots, n$$

can successively be calculated by

$$m^{(k+1)} = Am^{(k)} \tag{4.31}$$

with the $n \times n$ matrix

$$A = \begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & \text{\Large 0} & \\ & \text{\Large 0} & \ddots & \ddots & \\ & & & 1 & 1 \\ 1 & 1 & \cdots & 0 & 1 \end{pmatrix}.$$

The matrix A is nonnegative and its directed graph (see [54], p.20) is strongly connected, i.e., A is irreducible. The Perron-Frobenius theorem (see [54], p.30) implies that A has a positive eigenvalue γ_1 equal to its spectral radius. On the other hand, by a simple application of Theorem 2.9 in [54], A is also primitive. Thus, for the remaining eigenvalues $\gamma_2, \gamma_3, \ldots, \gamma_n$ of A, we have

$$\gamma_1 = \rho(A) > |\gamma_2| \geq |\gamma_3| \geq \cdots \geq |\gamma_n|. \tag{4.32}$$

Let

$$A^k = \left(a_{ij}^{(k)}\right), \quad k = 1, 2, \ldots,$$

denote the kth power of A. Since A is an irreducible matrix,

$$A^k > 0 \quad \text{for } k \geq k_0 \tag{4.33}$$

(see [54], p.41). It can be shown that

$$\lim_{k \to \infty} \frac{a_{ij}^{(k+1)}}{a_{ij}^{(k)}} = \gamma_1.$$

If $\epsilon > 0$ is given, then

$$a_{ij}^{(k+1)}/a_{ij}^{(k)} \geq \rho(A) - \epsilon \quad \text{for } k \geq k(\epsilon) \geq k_0,$$

or

$$a_{ij}^{(k+1)} \geq \alpha[\rho(A) - \epsilon], \quad i, \ j = 1, 2, \ldots, n,$$

where

$$\alpha = \min_{1 \leq i,j \leq n} a_{ij}^{(k)} > 0.$$

Therefore,

$$a_{ij}^{(k+2)} \geq a_{ij}^{(k+1)}[\rho(A) - \epsilon] \geq \alpha[\rho(A) - \epsilon]^2,$$

and, in general,

$$a_{ij}^{(k+r)} \geq \alpha[\rho(A) - \epsilon]^r, \quad i, \ j = 1, 2, \ldots, n, \quad r = 1, 2, \ldots. \tag{4.34}$$

Now, combining (4.31) and (4.34) into a single inequality

$$\begin{aligned} m^{(k+r)} &= A^{k+r} m^{(0)} = \left(\sum_{j=1}^{n} a_{ij}^{(k+r)} \right) \\ &\geq \left(n\alpha[\rho(A) - \epsilon]^r \right) e, \end{aligned}$$

where $e = (e_i)$, $e_i = 1$, $i = 1, 2, \ldots, n$, we obtain

$$\begin{aligned} \eta_i^{(k+r)} &\leq \eta^{m_i^{(k+r)}} \leq \eta^{n\alpha[\rho(A)-\epsilon]^r}; \\ \rho_i^{(k+r)} &\leq \eta^{m_i^{(k+r)}} \leq \eta^{n\alpha[\rho(A)-\epsilon]^r}, \end{aligned}$$

for $i = 1, 2, \ldots, n,\quad r = 1, 2, \ldots,\quad k \geq k(\epsilon) \geq k_0$, or

$$\begin{aligned} \varepsilon_i^{(k+r)} &\leq \frac{1}{\alpha}\eta^{n\alpha[\rho(A)-\epsilon]^r}; \\ \delta_i^{(k+r)} &\leq \frac{1}{\alpha}\eta^{n\alpha[\rho(A)-\epsilon]^r}. \end{aligned}$$

For

$$\varepsilon^{(k)} = \max_{1\leq i\leq n} \varepsilon_i^{(k)} \quad \text{and} \quad \delta^{(k)} = \max_{1\leq i\leq n} \delta_i^{(k)}$$

we also have

$$\begin{aligned} \varepsilon^{(k+r)} &\leq \frac{1}{\alpha}\eta^{n\alpha[\rho(A)-\epsilon]^r}; \\ \delta^{(k+r)} &\leq \frac{1}{\alpha}\eta^{n\alpha[\rho(A)-\epsilon]^r}. \end{aligned}$$

It follows that

$$\begin{aligned} \mathbf{R}_{\rho(A)-\epsilon}\{\varepsilon^{(k)}\} &= \lim_{r\to\infty} \sup \left[\varepsilon^{(k+r)}\right]^{1/[\rho(A)-\epsilon]^r} \\ &\leq \lim_{r\to\infty} \sup \left[\frac{1}{\alpha}\eta^{\alpha n[\rho(A)-\epsilon]^r}\right]^{1/[\rho(A)-\epsilon]^r} \\ &= \eta^{\alpha n} < 1. \end{aligned}$$

Similarly,

$$\mathbf{R}_{\rho(A)-\epsilon}\{\delta^{(k)}\} \leq \eta^{\alpha n} < 1.$$

Therefore,

$$O_{\mathbf{R}}((SSIMS), \lambda_i) \geq \rho(A) - \epsilon.$$

This inequality holds for all $\epsilon > 0$ and we immediately have

$$O_{\mathbf{R}}((SSIMS), \lambda_i) \;\geq\; \rho(A). \tag{4.35}$$

We now consider the characteristic polynomial $q_n(\gamma)$ of A:

$$q_n(\gamma) = (\gamma - 1)^n - (\gamma - 1) - 1.$$

If we set $\sigma = \gamma - 1$, then simply substituting this in the polynomial above yields

$$q_n(\sigma) = \sigma^n - \sigma - 1.$$

Since $q_n(1) = -1$ and $q_n(2) > 0$ for $n \geq 2$, there is a root σ_n with $1 < \sigma_n < 2$, and by Descartes' rule of signs, there can be no other positive root of $q_n(\sigma)$. Thus, for the spectral radius $\rho(A)$ of A we have

$$\rho(A) = 1 + \sigma_n,$$

and the combination with (4.35) gives

$$O_{\mathbf{R}}((SSIMS), \lambda_i) \geq 1 + \sigma_n$$

which completes the proof. □

Now consider the multiple-zero case. For the polynomial in (4.14), we define the following single-step iterations (denoted by *SSIMM*):

$$x_i^{(k+1)} = x_i^{(k)} + \sqrt[n_i]{\left| \frac{P(x_i^{(k)})}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)})^{n_j} \prod_{j=i+1}^{m}(x_i^{(k)} - y_j^{(k)})^{n_j}} \right|} \equiv C_{n_i}(x_i^{(k)}) \; ; \tag{4.36}$$

$$y_i^{(k+1)} = y_i^{(k)} - \sqrt[n_i]{\frac{P(y_i^{(k)})}{\prod_{j=1}^{i-1}(y_i^{(k)} - x_j^{(k+1)})^{n_j} \prod_{j=i+1}^{m}(y_i^{(k)} - y_j^{(k)})^{n_j}}} \equiv D_{n_i}(y_i^{(k)}) \; , \tag{4.37}$$

for $i = 1, \ldots, m$ and $k = 0, 1, 2, \ldots$.

As in Theorem 4.3.1 and Theorem 4.3.2 we can show the following two theorems.

Theorem 4.4.3 *For the polynomial (4.14), if the initial values* $\{x_i^{(0)}\}_{i=1}^m$ *and* $\{y_i^{(0)}\}_{i=1}^m$ *satisfy :*

$$x_1^{(0)} \leq \lambda_1 \leq y_1^{(0)} < x_2^{(0)} \leq \lambda_2 \leq y_2^{(0)} < \cdots < x_m^{(0)} \leq \lambda_m \leq y_m^{(0)}, \tag{4.38}$$

then

(a) *The sequence $\{x_i^{(k)}\}_{k=0}^{\infty}$ generated by (4.36) converges to λ_i monotonically. That is,*

$$x_i^{(0)} < x_i^{(1)} < \cdots < x_i^{(k)} \overset{k\to\infty}{\longrightarrow} \lambda_i, \quad i = 1, \ldots, m. \tag{4.39}$$

(b) *The sequence $\{y_i^{(k)}\}_{k=0}^{\infty}$ generated by (4.37) converges to λ_i monotonically. That is,*

$$\lambda_i \overset{k\to\infty}{\longleftarrow} y_i^{(k)} < y_i^{(k-1)} < \cdots < y_i^{(1)} < y_i^{(0)}, \quad i = 1, \ldots, m. \tag{4.40}$$

Theorem 4.4.4 *If $P(z) = \prod_{i=1}^{n}(z-\lambda_i)$ has a cluster of zeros $\{\lambda_i^{(1)}, \lambda_i^{(2)}, \ldots, \lambda_i^{(n_i)}\}_{i=1}^{m}$ with $\lambda_i^{(1)} \leq \lambda_i^{(2)} \leq \cdots \leq \lambda_i^{(n_i)} < \lambda_{i+1}^{(1)}$ and the initial values $\{x_i\}_{i=1}^{m}$ and $\{y_i\}_{i=1}^{m}$ satisfy*

$$\begin{aligned} x_1 &\leq \lambda_1^{(1)} \leq \cdots \leq \lambda_1^{(n_1)} \leq y_1 < x_2 \leq \lambda_2^{(1)} \leq \cdots \\ &\leq \lambda_2^{(n_2)} \leq y_2 < \cdots < x_m \leq \lambda_m^{(1)} \leq \cdots \leq \lambda_m^{(n_m)} \leq y_m \end{aligned} \tag{4.41}$$

where $\sum_{i=1}^{m} n_i = n$, then

$$x_i < C_1(x_i) < C_2(x_i) < \cdots < C_{n_i}(x_i) < \lambda_i^{(n_i)}, \quad i = 1, \ldots, m \tag{4.42}$$

and

$$\lambda_i^{(1)} < D_{n_i}(y_i) < D_{n_i-1}(y_i) < \cdots < D_1(y_i) < y_i, \quad i = 1, \ldots, m \tag{4.43}$$

where $C_{n_i}(x_i)$ and $D_{n_i}(y_i)$ are the same as in the iterations (4.36) and (4.37).

As in Theorem 4.4.2 we can show the following theorem.

Theorem 4.4.5 *Let $\sigma_m > 1$ be the unique positive root of $q_m(\sigma) = \sigma^m - \sigma - 1 = 0$. Then for the* **R**-*order of (SSIMM), we have*

$$O_{\mathbf{R}}((SSIMM), \lambda_i) \geq 1 + \sigma_m.$$

Proof Let $\alpha_{ij} = \frac{n_i}{n_j}$, $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i, j = 1, \ldots, n$. Then, by (4.36), we have

$$\begin{aligned} \varepsilon_i^{(k+1)} &= \varepsilon_i^{(k)} \left[1 - \frac{\prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j)^{\alpha_{ij}} \prod_{j=i+1}^{m}(\lambda_j - x_i^{(k)})^{\alpha_{ij}}}{\prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)})^{\alpha_{ij}} \prod_{j=i+1}^{m}(y_j^{(k)} - x_i^{(k)})^{\alpha_{ij}}} \right] \\ &= \varepsilon_i^{(k)} \frac{N_3}{D_3} \end{aligned}$$

where

$$D_3 = \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)})^{\alpha_{ij}} \prod_{j=i+1}^{m} (y_j^{(k)} - x_i^{(k)})^{\alpha_{ij}}$$

and

$$\begin{aligned} N_3 &= \prod_{j=1}^{i-1}(x_i^{(k)} - x_j^{(k+1)})^{\alpha_{ij}} \prod_{j=i+1}^{m} (y_j^{(k)} - x_i^{(k)})^{\alpha_{ij}} \\ &\quad - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j)^{\alpha_{ij}} \prod_{j=i+1}^{m} (\lambda_j - x_i^{(k)})^{\alpha_{ij}} \\ &= \prod_{j=1}^{i-1}[(x_i^{(k)} - \lambda_j) + \varepsilon_j]^{\alpha_{ij}} \prod_{j=i+1}^{m} [(\lambda_j - x_i^{(k)}) + \delta_j^{(k)}]^{\alpha_{ij}} \\ &\quad - \prod_{j=1}^{i-1}(x_i^{(k)} - \lambda_j)^{\alpha_{ij}} \prod_{j=i+1}^{m} (\lambda_j - x_i^{(k)})^{\alpha_{ij}}. \end{aligned}$$

It follows from the proof of Theorem 4.4.2 that $\exists c_3 > 0$ such that

$$\begin{aligned} \varepsilon_i^{(k+1)} &\le \varepsilon_i^{(k)} \left|\frac{N_3}{D_3}\right| \\ &\le C_3\varepsilon_i^{(k)} \left[\sum_{j=1}^{i-1} \varepsilon_j^{(k+1)} + \sum_{j=i+1}^{m} \delta_j^{(k)}\right] \end{aligned}$$

and $\exists C_4 > 0$ such that

$$\delta_i^{(k+1)} \le C_4\delta_i^{(k)} \left[\sum_{j=1}^{i-1} \varepsilon_j^{(k+1)} + \sum_{j=i+1}^{m} \delta_j^{(k)}\right].$$

Then, by applying the same steps as in Theorem 4.4.2, the assertion of the theorem follows. □

4.5 Numerical results.

Numerical results presented in this section illustrate the monotonic convergence of our new algorithms. In Table 4.1, iterations in (4.3) and (4.4) are applied on the polynomial $P(z) = z(z^2 - 1)(z^2 - 4)(z^2 - 9)$.

k	$x_1^{(k)}$	$y_1^{(k)}$	$x_2^{(k)}$	$y_2^{(k)}$	$x_3^{(k)}$	$y_3^{(k)}$
0	-3.5	-2.5	-2.5	-1.5	-1.5	-0.5
1	-3.29052734	-2.61279296	-2.38720703	-1.59228515	-1.40771484	-0.58544921
2	-3.15548526	-2.73862000	-2.25311124	-1.70800722	-1.29167636	-0.69429426
3	-3.06939121	-2.86149374	-2.12835706	-1.83431295	-1.16506581	-0.81893982
4	-3.02104604	-2.95343246	-2.04253799	-1.93972918	-1.06027549	-0.93056202
5	-3.00276308	-2.99364358	-2.00607224	-1.99101009	-1.00924659	-0.98903675
6	-3.00006028	-2.99986053	-2.00014563	-1.99978274	-1.00023501	-0.99971974
7	-3.00000003	-2.99999992	-2.00000008	-1.99999987	-1.00000015	-0.99999982
8	-3.00000000	-3.00000000	-2.00000000	-2.00000000	-1.00000000	-1.00000000

$x_4^{(k)}$	$y_4^{(k)}$	$x_5^{(k)}$	$y_5^{(k)}$	$x_6^{(k)}$	$y_6^{(k)}$	$x_7^{(k)}$	$y_7^{(k)}$
-0.5	0.5	0.5	1.5	1.5	2.5	2.5	3.5
-0.41455078	0.41455078	0.58544921	1.40771484	1.59228515	2.38720703	2.61279296	3.29052734
-0.30569628	0.30569628	0.69429426	1.29167636	1.70800722	2.25311124	2.73862000	3.15548526
-0.18113671	0.18113671	0.81893982	1.16506581	1.83431295	2.12835706	2.86149374	3.06939121
-0.06969138	0.06969138	0.93056202	1.06027549	1.93972918	2.04253799	2.95343246	3.02104604
-0.01111398	0.01111398	0.98903675	1.00924659	1.99101009	2.00607224	2.99364358	3.00276308
-0.00028864	0.00028864	0.99971974	1.00023501	1.99978274	2.00014563	2.99986053	3.00006028
-0.00000018	0.00000018	0.99999981	1.00000015	1.99999986	2.00000008	2.99999992	3.00000003
0.00000000	0.00000000	1.00000000	1.00000000	2.00000000	2.00000000	3.00000000	3.00000000

Table 4.1: Numerical solutions by the new methods on $P(z) = z(z^2-1)(z^2-4)(z^2-9)$

Chapter 5

Solving Symmetric Tridiagonal Eigenvalue Problems

5.1 Introduction

The homotopy continuation algorithms for eigenvalue problems were mainly developed by T. Y. Li and his former students (see [15], [36] and [35]). Finding the eigenvalues of an $n \times n$ symmetric tridiagonal matrix A is equivalent to solving $\det[A - \lambda I] = 0$ which is a polynomial of a single variable λ with degree n having only real zeros.

In this chapter, we present an algorithm for the eigenvalue problem of symmetric tridiagonal matrices. Our algorithm is parallel in nature. We first propose the following new concurrent iterations with cubically convergent rate:

$$\begin{aligned} x_i^{(k+1)} &= x_i^{(k)} - \frac{n_i P(x_i^{(k)})}{P'(x_i^{(k)}) - P(x_i^{(k)}) \sum_{j=1, j\neq i}^{m} \frac{n_j}{x_i^{(k)} - y_j^{(k)}}}, \quad k = 0, 1, 2, \ldots, \\ y_i^{(k+1)} &= y_i^{(k)} - \frac{n_i P(y_i^{(k)})}{P'(y_i^{(k)}) - P(y_i^{(k)}) \sum_{j=1, j\neq i}^{m} \frac{n_j}{y_i^{(k)} - x_j^{(k)}}}, \quad k = 0, 1, 2, \ldots, \end{aligned}$$

where $P(\lambda) = \prod_{j=1}^{m} (\lambda - \lambda_j)^{n_j}$, $\sum_{j=1}^{m} n_j = n$ and $i = 1, 2, \ldots, m$. The monotonic convergence of these concurrent iterations for real multiple zeros of $P(\lambda)$ will be given in Section 5.2 with non-overshoot properties for clusters of real zeros of $P(\lambda)$. Moreover, these iterations without second derivative evaluations are cubically convergent. To solve the eigenvalue problems, we first employ the Sturm sequence to detect the multiplicities of the eigenvalues. Then, our new concurrent iterative methods are used to extract the eigenvalues iteratively by using eigenvalues of two smaller matrices as the initial approximations. The numerical results seem remarkable.

5.2 A new method with cubic convergence rate and monotonic convergence

Consider a monic polynomial of degree $n \geq 3$,

$$P(\lambda) = \prod_{i=1}^{n} (\lambda - \lambda_i), \tag{5.1}$$

with simple real zeros

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n. \tag{5.2}$$

First, we construct a pair of sequences $\{x_i^{(k)}\}_{i=1}^n$ and $\{y_i^{(k)}\}_{i=1}^n$ by the following iterations with a pair of initial values $\{x_i^{(0)}\}_{i=1}^n$ and $\{y_i^{(0)}\}_{i=1}^n$:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{P(x_i^{(k)})}{P'(x_i^{(k)}) - P(x_i^{(k)}) \sum_{j=1, j\neq i}^{n} \frac{1}{x_i^{(k)} - y_j^{(k)}}}, \quad i = 1, \ldots, n, \tag{5.3}$$

$$y_i^{(k+1)} = y_i^{(k)} - \frac{P(y_i^{(k)})}{P'(y_i^{(k)}) - P(y_i^{(k)}) \sum_{j=1, j\neq i}^{n} \frac{1}{y_i^{(k)} - x_j^{(k)}}}, \quad i = 1, \ldots, n, \tag{5.4}$$

for $k = 0, 1, 2, \ldots$. These two sequences can be used to approximate all zeros of the polynomial $P(\lambda)$ simultaneously. The convergence rates of the iterations in (5.3) and (5.4) are both cubic for a simple zero. However, they converge only linearly for multiple zeros. For a monic polynomial of degree $n \geq 3$,

$$P(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_i)^{n_i}, \tag{5.5}$$

with multiple zeros

$$\lambda_1 < \lambda_2 < \cdots < \lambda_m \text{ and } \sum_{i=1}^{m} n_i = n, \tag{5.6}$$

we can modify the iterations in (5.3) and (5.4) as follows:

$$x_i^{(k+1)} = x_i^{(k)} - \frac{n_i P(x_i^{(k)})}{P'(x_i^{(k)}) - P(x_i^{(k)}) \sum_{j=1,j\neq i}^{m} \frac{n_j}{x_i^{(k)} - y_j^{(k)}}} \equiv A_{n_i}(x_i^{(k)}) \,, \qquad (5.7)$$

$$y_i^{(k+1)} = y_i^{(k)} - \frac{n_i P(y_i^{(k)})}{P'(y_i^{(k)}) - P(y_i^{(k)}) \sum_{j=1,j\neq i}^{m} \frac{n_j}{y_i^{(k)} - x_j^{(k)}}} \equiv B_{n_i}(y_i^{(k)}) \qquad (5.8)$$

for $i = 1, \ldots, m$ and $k = 0, 1, 2, \ldots$, and with the initial values $\{x_i^{(0)}\}_{i=1}^m$ and $\{y_i^{(0)}\}_{i=1}^m$.

The above iterations in (5.7) and (5.8) are obtained by applying one step of modified Newton's method to the following rational functions:

$$R_i^{(k)}(x) = \frac{P(x)}{\prod_{j=1,j\neq i}^{m}(x - y_j^{(k)})^{n_j}},$$

$$S_i^{(k)}(y) = \frac{P(y)}{\prod_{j=1,j\neq i}^{m}(y - x_j^{(k)})^{n_j}}.$$

In the following two theorems, we will show that the iterations defined in (5.7) and (5.8) are monotonically and cubically convergent for multiple zeros.

Theorem 5.2.1 *For the polynomial $P(\lambda)$ in (5.5), if the initial values $\{x_i^{(0)}\}_{i=1}^m$ and $\{y_i^{(0)}\}_{i=1}^m$ satisfy*

$$x_1^{(0)} \le \lambda_1 \le y_1^{(0)} < x_2^{(0)} \le \lambda_2 \le y_2^{(0)} < \cdots < x_m^{(0)} \le \lambda_m \le y_m^{(0)}, \qquad (5.9)$$

then

(a) *The sequence $\{x_i^{(k)}\}_{k=0}^{\infty}$ generated by (5.7) converges to λ_i monotonically. That is,*

$$x_i^{(0)} < x_i^{(1)} < \cdots < x_i^{(k)} \stackrel{k\to\infty}{\longrightarrow} \lambda_i, \quad i = 1, \ldots, m. \qquad (5.10)$$

(b) *The sequence $\{y_i^{(k)}\}_{k=0}^{\infty}$ generated by (5.8) converges to λ_i monotonically. That is,*

$$\lambda_i \stackrel{k\to\infty}{\longleftarrow} y_i^{(k)} < y_i^{(k-1)} < \cdots < y_i^{(1)} < y_i^{(0)}, \quad i = 1, \ldots, m. \qquad (5.11)$$

Proof Let $\varepsilon_i^{(k)} = \lambda_i - x_i^{(k)}$ and $\delta_i^{(k)} = y_i^{(k)} - \lambda_i$ for $i = 1, \ldots, m$. We have, by (5.7),

$$\begin{aligned}
\varepsilon_i^{(k+1)} &= \varepsilon_i^{(k)} + \frac{n_i}{\frac{P'(x_i^{(k)})}{P(x_i^{(k)})} - \sum_{j=1,j\neq i}^{m} \frac{n_j}{x_i^{(k)} - y_j^{(k)}}} \\
&= \varepsilon_i^{(k)} + \frac{n_i}{\sum_{j=1,j\neq i}^{m} \left(\frac{n_j}{x_i^{(k)} - \lambda_j} - \frac{n_j}{x_i^{(k)} - y_j^{(k)}} \right) + \frac{n_i}{x_i^{(k)} - \lambda_i}} \\
&= \varepsilon_i^{(k)} - \frac{(\lambda_i - x_i^{(k)}) n_i}{(x_i^{(k)} - \lambda_i) \sum_{j=1,j\neq i}^{m} \left(\frac{n_j}{x_i^{(k)} - \lambda_j} - \frac{n_j}{x_i^{(k)} - y_j^{(k)}} \right) + n_i} \\
&= \varepsilon_i^{(k)} \left(1 - \frac{n_i}{\Psi_i^{(k)} + n_i} \right) \\
&= \varepsilon_i^{(k)} \cdot \frac{\Psi_i^{(k)}}{\Psi_i^{(k)} + n_i} \qquad (5.12)
\end{aligned}$$

where

$$\Psi_i^{(k)} = (x_i^{(k)} - \lambda_i) \sum_{j=1,j\neq i}^{m} \left(\frac{n_j}{x_i^{(k)} - \lambda_j} - \frac{n_j}{x_i^{(k)} - y_j^{(k)}} \right)$$

for $i = 1, 2, \ldots, m$.

By (5.9), we know $\Psi_i^{(0)} > 0$. Thus,

$$0 < \varepsilon_i^{(1)} = \varepsilon_i^{(0)} \cdot \frac{\Psi_i^{(0)}}{\Psi_i^{(0)} + n_i} < \varepsilon_i^{(0)}.$$

That is,

$$x_i^{(0)} < x_i^{(1)} < \lambda_i.$$

Similarly, by (5.8), we have

$$\begin{aligned}
\delta_i^{(k+1)} &= \delta_i^{(k)} - \frac{n_i}{\frac{P'(y_i^{(k)})}{P(y_i^{(k)})} - \sum_{j=1,j\neq i}^{m} \frac{n_j}{y_i^{(k)} - x_j^{(k)}}} \\
&= \delta_i^{(k)} - \frac{n_i}{\sum_{j=1,j\neq i}^{m} \left(\frac{n_j}{y_i^{(k)} - \lambda_j} - \frac{n_j}{y_i^{(k)} - x_j^{(k)}} \right) + \frac{n_i}{y_i^{(k)} - \lambda_i}} \\
&= \delta_i^{(k)} - \frac{(y_i^{(k)} - \lambda_i) n_i}{(y_i^{(k)} - \lambda_i) \sum_{j=1,j\neq i}^{m} \left(\frac{n_j}{y_i^{(k)} - \lambda_j} - \frac{n_j}{y_i^{(k)} - x_j^{(k)}} \right) + n_i} \\
&= \delta_i^{(k)} \left(1 - \frac{n_i}{\Phi_i^{(k)} + n_i} \right) \\
&= \delta_i^{(k)} \cdot \frac{\Phi_i^{(k)}}{\Phi_i^{(k)} + n_i} \qquad (5.13)
\end{aligned}$$

where

$$\Phi_i^{(k)} = (y_i^{(k)} - \lambda_i) \sum_{j=1, j\neq i}^{m} \left(\frac{n_j}{y_i^{(k)} - \lambda_j} - \frac{n_j}{y_i^{(k)} - x_j^{(k)}} \right)$$

for $i = 1, 2, \ldots, m$.

By (5.9), we know $\Phi_i^{(0)} > 0$. Thus,

$$0 < \delta_i^{(1)} = \delta_i^{(0)} \cdot \frac{\Phi_i^{(0)}}{\Phi_i^{(0)} + n_i} < \delta_i^{(0)}.$$

That is,

$$\lambda_i < y_i^{(1)} < y_i^{(0)}.$$

Therefore, by mathematical induction, (5.10) and (5.11) can easily be obtained from (5.12) and (5.13). □

Theorem 5.2.2 *The convergence rates of the modified iterations in (5.7) and (5.8) are both cubic for an n_i-fold zeros λ_i of $P(\lambda)$.*

Proof It follows from the proof of the above theorem that

$$\varepsilon_i^{(k+1)} = \varepsilon_i^{(k)} \cdot \frac{\Psi_i^{(k)}}{\Psi_i^{(k)} + n_i}$$

where

$$\begin{aligned} \Psi_i^{(k)} &= (x_i^{(k)} - \lambda_i) \sum_{j=1, j\neq i}^{m} \left(\frac{n_j}{x_i^{(k)} - \lambda_j} - \frac{n_j}{x_i^{(k)} - y_j^{(k)}} \right) \\ &= (x_i^{(k)} - \lambda_i) \sum_{j=1, j\neq i}^{m} n_j \frac{(\lambda_j - y_j^{(k)})}{(x_i^{(k)} - \lambda_j)(x_i^{(k)} - y_j^{(k)})}. \end{aligned}$$

Let $e^{(k)} = \max_{i=1,\ldots,m}(|\varepsilon_i^{(k)}|, |\delta_i^{(k)}|)$. Then,

$$\varepsilon_i^{(k+1)} \sim (e^{(k)})^3.$$

Similarly, we have

$$\delta_i^{(k+1)} \sim (e^{(k)})^3.$$

□

Remark 5.2.1 *The monotonic convergence and the cubic convergence rates of the iterations in (5.3) and (5.4) can easily be obtained from the above two theorems by taking all* $n_i = 1$.

In the next theorem we show that the modified iterations in (5.7) and (5.8) have non-overshoot properties for polynomials with clusters of zeros. Namely, they never skip over n_i zeros of a polynomial. The non-overshoot properties are important when they are applied to find the eigenvalues of a symmetric matrix.

Theorem 5.2.3 *If* $P(\lambda) = \prod_{i=1}^{n}(\lambda - \lambda_i)$ *has a cluster of zeros* $\{\lambda_i^{(1)}, \lambda_i^{(2)}, \ldots, \lambda_i^{(n_i)}\}_{i=1}^{m}$ *with* $\lambda_i^{(1)} \le \lambda_i^{(2)} \le \cdots \le \lambda_i^{(n_i)} < \lambda_{i+1}^{(1)}$ *and the initial values* $\{x_i\}_{i=1}^{m}$ *and* $\{y_i\}_{i=1}^{m}$ *satisfy*

$$\begin{aligned} x_1 &\le \lambda_1^{(1)} \le \cdots \le \lambda_1^{(n_1)} \le y_1 < x_2 \le \lambda_2^{(1)} \le \cdots \\ &\le \lambda_2^{(n_2)} \le y_2 < \cdots < x_m \le \lambda_m^{(1)} \le \cdots \le \lambda_m^{(n_m)} \le y_m \end{aligned} \tag{5.14}$$

where $\sum_{i=1}^{m} n_i = n$, *then*

$$x_i < A_1(x_i) < A_2(x_i) < \cdots < A_{n_i}(x_i) < \lambda_i^{(n_i)}, \quad i = 1, \ldots, m \tag{5.15}$$

and

$$\lambda_i^{(1)} < B_{n_i}(y_i) < B_{n_i-1}(y_i) < \cdots < B_1(y_i) < y_i, \quad i = 1, \ldots, m \tag{5.16}$$

where $A_{n_i}(x_i)$ *and* $B_{n_i}(y_i)$ *are the same as in the iterations (5.7) and (5.8).*

Proof By (5.7), (5.8) and (5.14), we have

$$\begin{aligned} \lambda_i^{(n_i)} - A_{n_i}(x_i) &> (\lambda_i^{(n_i)} - x_i) + \frac{n_i}{\sum_{j=1, j\ne i}^{m}\left(\frac{n_j}{x_i - \lambda_j^{(n_j)}} - \frac{n_j}{x_i - y_j}\right) + \frac{n_i}{x_i - \lambda_i^{(n_i)}}} \\ &= (\lambda_i^{(n_i)} - x_i)\left(1 - \frac{n_i}{\Psi_i^{(n_i)} + n_i}\right) \\ &> 0; \end{aligned}$$

$$\begin{aligned} B_{n_i}(y_i) - \lambda_i^{(1)} &> (y_i - \lambda_i^{(1)}) - \frac{n_i}{\sum_{j=1, j\ne i}^{m}\left(\frac{n_j}{y_i - \lambda_j^{(1)}} - \frac{n_j}{y_i - x_j}\right) + \frac{n_i}{y_i - \lambda_i^{(1)}}} \\ &= (y_i - \lambda_i^{(1)})\left(1 - \frac{n_i}{\Phi_i^{(1)} + n_i}\right) \\ &> 0. \end{aligned}$$

and it is easy to check that:

$$\frac{\partial A_{n_i}(x_i)}{\partial n_i} > 0 \text{ and } \frac{\partial B_{n_i}(y_i)}{\partial n_i} < 0.$$

Therefore, (5.15) and (5.16) follow. □

5.3 Evaluation of the logarithmic derivative P'/P of the determinant

Let T be a symmetric tridiagonal matrix of the form

$$T = [\beta_{i-1}, \alpha_i, \beta_i] = \begin{pmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & 0 & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & 0 & & \beta_{n-2} & \alpha_{n-1} & \beta_{n-1} \\ & & & & \beta_{n-1} & \alpha_n \end{pmatrix}. \tag{5.17}$$

We may assume, without loss of generality, that T is unreduced; that is, $\beta_j \neq 0$, $j = 1, \ldots, n-1$. For an unreduced T, the characteristic polynomial

$$P(\lambda) \equiv \det(T - \lambda I) \tag{5.18}$$

has only *real* and *simple* zeros ([56], p300). In order to use our iterations developed in previous section for finding zeros of $P(\lambda)$, or the eigenvalues of T, it is necessary to evaluate P and P' efficiently with satisfactory accuracy in the first place.

It is well known that the characteristic polynomial $P(\lambda) \equiv \det(T - \lambda I)$ of $T = [\beta_{i-1}, \alpha_i, \beta_i]$ and its derivative with respect to λ can be evaluated by three-term recurrences ([56], p423):

$$\begin{cases} \rho_0 = 1, \quad \rho_1 = \alpha_1 - \lambda \\ \rho_i = (\alpha_i - \lambda)\rho_{i-1} - \beta_{i-1}^2 \rho_{i-2}, \quad i = 2, 3, \ldots, n \end{cases} \tag{5.19}$$

$$\begin{cases} \rho'_0 = 0, \quad \rho'_1 = -1 \\ \rho'_i = (\alpha_i - \lambda)\rho'_{i-1} - \rho_{i-1} - \beta_{i-1}^2 \rho'_{i-2}, \quad i = 2, 3, \ldots, n \end{cases} \tag{5.20}$$

and

$$P(\lambda) = \rho_n, \qquad P'(\lambda) = \rho_n'.$$

However, these recurrences may suffer from a severe underflow-overflow problem and require constant testing and scaling. The following modified recurrence equations avoid the above problem subtly by computing the quotient $q(\lambda) = P'(\lambda)/P(\lambda)$ directly [36]. After all, only $q(\lambda)$ is really needed in the formulae (5.3) and (5.4), or in (5.7) and (5.8). Let

$$\xi_i = \frac{\rho_i}{\rho_{i-1}}, \qquad \eta_i = -\frac{\rho_i'}{\rho_i}$$

$$\begin{cases} \xi_1 = \alpha_1 - \lambda, \\ \xi_i = \alpha_i - \lambda - \dfrac{\beta_{i-1}^2}{\xi_{i-1}}, \quad i = 2, 3, \ldots, n. \end{cases} \tag{5.21}$$

$$\begin{cases} \eta_0 = 0, \quad \eta_1 = \dfrac{1}{\xi_1} \\ \eta_i = \dfrac{1}{\xi_i} \left[(\alpha_i - \lambda)\eta_{i-1} + 1 - \left(\dfrac{\beta_{i-1}^2}{\xi_{i-1}} \right) \eta_{i-2} \right], \quad i = 2, 3, \ldots, n \end{cases} \tag{5.22}$$

and

$$-\frac{P'(\lambda)}{P(\lambda)} = \eta_n.$$

To prevent the algorithm from breaking down when $\xi_i = 0$ for some $1 \le i \le n$, an extra check is provided:

- If $\xi_1 = 0$ (i.e., $\alpha_1 = \lambda$), set $\xi_1 = \beta_1^2 \varepsilon^2$;
- If $\xi_i = 0$, $i > 1$, set $\xi_i = \dfrac{\beta_{i-1}^2 \varepsilon^2}{\xi_{i-1}}$;

where ε is the machine precision. A determinant evaluation subroutine DETEVL (see Figure 5.1) is formulated according to the recurrences (5.21) and (5.22). When ξ_i, $i = 1, \ldots, n$ are known, the Sturm sequence is available ([40], p47). Thus, as a by-product, DETEVL also evaluates the number of eigenvalues of T which are less than λ. This number is denoted by $\kappa(\lambda)$.

A backward error analysis in [36] shows that the accuracy of this algorithm at λ_i can be

$$\frac{5\varepsilon}{2} \max_j (|\beta_j| + |\beta_{j+1}|) + |\lambda_i| \varepsilon. \tag{5.23}$$

5.4 The split-merge process

Let

$$\lambda_1 < \lambda_2 < \cdots < \lambda_n$$

be the zeros of P in (5.18). Using the iterations in (5.3) and (5.4) to approximate any λ_i, $i = 1, 2, \ldots, n$, it is essential to provide a pair of starting points $x_i^{(0)}$ and $y_i^{(0)}$, being $x_i^{(0)} < \lambda_i < y_i^{(0)}$. For this purpose, we *split* the matrix T into

$$\hat{T} = \begin{pmatrix} T_0 & 0 \\ 0 & T_1 \end{pmatrix} \tag{5.24}$$

where

$$T_0 = \begin{pmatrix} \alpha_1 & \beta_1 & & \\ \beta_1 & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{k-1} \\ & & \beta_{k-1} & \alpha_k - \beta_k \end{pmatrix}, \quad T_1 = \begin{pmatrix} \alpha_{k+1} - \beta_k & \beta_{k+1} & & \\ \beta_{k+1} & \ddots & \ddots & \\ & \ddots & \ddots & \beta_{n-1} \\ & & \beta_{n-1} & \alpha_n \end{pmatrix}. \tag{5.25}$$

Obviously, the eigenvalues of $\hat{T}$ consist of eigenvalues of T_0 and T_1. Without loss of generality, we may assume $\beta_i > 0$, for all $i = 1, 2, \ldots, n-1$, since in (5.19)–(5.22), β_i's always appear in their square form. The following interlacing property for this rank-one tearing is important to our algorithm.

Theorem 5.4.1 *Let $\lambda_1 < \lambda_2 < \cdots < \lambda_n$ and $\hat{\lambda}_1 < \hat{\lambda}_2 < \cdots < \hat{\lambda}_n$ be eigenvalues of T and $\hat{T}$ respectively. Then*

$$\hat{\lambda}_1 \le \lambda_1 \le \hat{\lambda}_2 \le \lambda_2 \le \cdots \le \hat{\lambda}_n \le \lambda_n < \hat{\lambda}_{n+1},$$

with the convention $\hat{\lambda}_{n+1} = \hat{\lambda}_n + 2\beta_k$.

Proof See [23, Theorem 8.6.2, p462].

The eigenvalues of $\hat{T}$ will be used critically to approximate the eigenvalues of T by our new iterations in (5.3) and (5.4). We will call this procedure, splitting T into T_0 and T_1 of $\hat{T}$ and using eigenvalues of $\hat{T}$, consisting of eigenvalues of T_0 and

T_1, to approximate eigenvalues of T, the *split-merge* process, similar to Cuppen's divide-and-conquer strategy [9], of course.

The eigenvalues of $\hat{T}$ in (5.24) consist of eigenvalues of T_0 and T_1 in (5.25). To find eigenvalues of T_0 and T_1, the split-merge process described above may be applied to T_0 and T_1 again. Indeed, the splitting process should be applied to T recursively (See Figure 5.2) until 2×2 and 1×1 matrices are reached.

After T is well split into a tree structure as shown in Figure 5.2, the merging process in the reverse direction from 2×2 and 1×1 matrices can be started. More specifically, let T_σ be split into $T_{\sigma 0}$ and $T_{\sigma 1}$. Let $\hat{\lambda}_1^\sigma, \cdots \hat{\lambda}_m^\sigma$ be eigenvalues of $\hat{T}_\sigma = \begin{pmatrix} T_{\sigma 0} & 0 \\ 0 & T_{\sigma 1} \end{pmatrix}$ in ascending order. Then the iterations (5.3) and (5.4) is applied to the polynomial equation

$$P_\sigma(\lambda) \equiv \det[T_\sigma - \lambda I] = 0$$

from $\hat{\lambda}_i^{(\sigma)}$ and $\hat{\lambda}_{i+1}^{(\sigma)}$ to obtain the corresponding eigenvalue λ_i^σ, $i = 1, 2, \ldots, m$, by the merging process described above. This process is continued until T_0 and T_1 are merged into T. That is, in the final step all the eigenvalues of T are obtained by applying our new iterations to $P(\lambda) = \det(T - \lambda I)$ from eigenvalues of T_0 and T_1.

5.5 Cluster spectrum

By Theorem 5.4.1, $\lambda_i \in (\hat{\lambda}_i, \hat{\lambda}_{i+1})$ for each $i = 1, \ldots, n$ with the convention $\hat{\lambda}_{n+1} = \hat{\lambda}_n + 2\beta_k$. If $\hat{\lambda}_{i+1} - \hat{\lambda}_i$ is less than the error tolerance, then either $\hat{\lambda}_i$ or $\hat{\lambda}_{i+1}$ can be accepted as λ_i. In general, if $\hat{T}$ has a cluster of $r+1$ very close eigenvalues, for instance, $\hat{\lambda}_{j+r} - \hat{\lambda}_j$ is less than the error tolerance for certain $1 \le j \le n - r$, then r eigenvalues $\lambda_j, \lambda_{j+1}, \ldots, \lambda_{j+r-1}$ of T can be obtained free of computations. They can be set to any one of $\hat{\lambda}_j, \ldots, \hat{\lambda}_{j+r}$.

Since the matrix T in (5.17) is unreduced, its eigenvalues are all simple. Therefore, using all $n_i = 1$ in the new iterations given in (5.7) and (5.8) seems appropriate in all cases to obtain ultimate super-linear convergence with cubic convergence rate.

However, in some occasions, there may exist a group of $r > 1$ eigenvalues of T, say,

$$\lambda_{i+1} < \lambda_{i+2} < \cdots < \lambda_{i+r}$$

which are relatively close to each other, comparing to the distance from the two starting points, say $x_{i+1}^{(0)} < \lambda_{i+1} < y_{i+1}^{(0)}$. Some of them may even be numerically indistinguishable. Numerical evidence shows that to reach λ_{i+1} from $x_{i+1}^{(0)}$ and $y_{i+1}^{(0)}$, it may take many steps of the new iterations given in (5.3) and (5.4) before showing super-linear convergence. In this situation, the new iterations given in (5.7) and (5.8) with $n_{i+1} = r$ may be used to speed up the convergence.

5.6 Stopping criteria and three iterative schemes

The following stopping criterion was suggested by Kahan [30]:

$$|x^{(k+1)} - x^{(k)}|^2 \leq (|x^{(k)} - x^{(k-1)}| - |x^{(k+1)} - x^{(k)}|)\tau \tag{5.26}$$

where τ is the error tolerance. This criterion is based on the following observations. Let

$$q_k = \left| \frac{x^{(k+1)} - x^{(k)}}{x^{(k)} - x^{(k-1)}} \right|,$$

then as $\{x^{(k)}\}_{k=1}^{\infty}$ converges to λ when $k \to \infty$, q_k is normally decreasing. Thus

$$\begin{aligned} |\lambda - x^{(k+1)}| &= \left| \sum_{i=0}^{\infty} (x^{(k+2+i)} - x^{(k+1+i)}) \right| \leq \sum_{i=0}^{\infty} |x^{(k+2+i)} - x^{(k+1+i)}| \\ &\leq |x^{(k+1)} - x^{(k)}| \sum_{i=1}^{\infty} q_k^i = \frac{q_k |x^{(k+1)} - x^{(k)}|}{1 - q_k} \\ &= \frac{|x^{(k+1)} - x^{(k)}|^2}{|x^{(k)} - x^{(k-1)}| - |x^{(k+1)} - x^{(k)}|}. \end{aligned}$$

From (5.23), an obvious error tolerance at λ_i can be chosen as

$$\tau = \frac{5\varepsilon}{2} \max_j (|\beta_j| + |\beta_{j+1}|) + \left| \frac{x^{(k+1)} + y^{(k+1)}}{2} \right| \varepsilon.$$

We use the following function for the following three Iterative Schemes:

$TEST(a, b) =' TRUE'$ if $|a - b| < \tau$ or $|a - b| < \varepsilon |a|$ where ε is the machine precision, $TEST(a, b) =' FALSE'$ otherwise.

In the following we present three Iterative Schemes for the merge process to compute an admissible set of points $\{\lambda_1^-, \lambda_1^+, \ldots, \lambda_n^-, \lambda_n^+\}$ for the set of eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$.

Iterative Scheme 1.

$S = \{1, \ldots, n\}, k = 1$

while $S \neq 0$ **do**

for $i \in S$ **do**

if $TEST(x_i^k, y_i^k) =' FALSE'$ **then** apply (5.3) and (5.4)

else $S = S - \{i\}$

enddo

for $i \notin S$ **do**

$x_i^{(k+1)} = x_i^{(k)},\ y_i^{(k+1)} = y_i^{(k)}.$

enddo

$k = k + 1$

enddo

for $i \in \{1, \ldots, n\}$ **do**

$\lambda_i^- = x_i^{(k)}, \lambda_i^+ = y_i^{(k)}$

enddo

The following Iterative Scheme 2 is the first modification of Iterative Scheme 1. For $i = 1, \ldots, n$, consider the midpoint $c_i^{(k)} = (x_i^{(k)} + y_i^{(k)})/2$ of the interval $[x_i^{(k)}, y_i^{(k)}]$. By applying the Sturm sequence at $c_i^{(k)}$, we may detect if the eigenvalue λ_i is smaller than or bigger than $c_i^{(k)}$. If $\lambda_i < c_i^{(k)}$, then we apply (5.3). Otherwise, we apply (5.4). We terminate the iteration if $TEST(x_i^{(k)}, y_i^{(k)}) =' TRUE'$. With this modification one of the two iterations (5.3) and (5.4) is replaced by the cheaper Sturm sequence computation.

Iterative Scheme 2.

$S = \{1, \ldots, n\}, k = 1$

while $S \neq 0$ **do**

for $i \in S$ **do**

$c_i^{(k)} = (x_i^{(k)} + y_i^{(k)})/2$

if $TEST(x_i^{(k)}, y_i^{(k)}) =' FALSE'$

then if $c_i^{(k)} > \lambda_i$

then $x_i^{(k+1)} = x_i^{(k)}, y_i^{(k+1)} = c_i^{(k)} - \frac{1}{P'(c_i^{(k)})/P(c_i^{(k)}) - \sum_{j=1, j\neq i}^{n} \frac{1}{c_i^{(k)} - x_j^{(k)}}}$

else $y_i^{(k+1)} = y_i^{(k)}, x_i^{(k+1)} = c_i^{(k)} - \frac{1}{P'(c_i^{(k)})/P(c_i^{(k)}) - \sum_{j=1, j\neq i}^{n} \frac{1}{c_i^{(k)} - y_j^{(k)}}}$

else $S = S - \{i\}$

enddo

for $i \notin S$ **do**

$x_i^{(k+1)} = x_i^{(k)},\ y_i^{(k+1)} = y_i^{(k)}$

enddo

$k = k + 1$

enddo

for $i \in \{1, \ldots, n\}$ **do**

$\lambda_i^- = x_i^{(k)}, \lambda_i^+ = y_i^{(k)}$

enddo

In the second modification we avoid using Sturm sequences at any iterations. For the first iteration, once we have detected which semi-interval λ_i belongs to, we keep applying (5.3) or (5.4) for all the iterations until either $TEST(x_i^{(k+1)}, x_i^{(k)}) =' TRUE'$ or $TEST(y_i^{(k+1)}, y_i^{(k)}) =' TRUE'$. In the first case we set $\lambda_i^- = x_i^{(k+1)}, \lambda_i^+ = \max\{x_i^{(k+1)}(1 + \varepsilon\, \text{sign}(x_i^{(k+1)})), x_i^{(k+1)} + \tau\}$, in the second case $\lambda_i^+ = y_i^{(k+1)}, \lambda_i^- = \max\{y_i^{(k+1)}(1 - \varepsilon\, \text{sign}(y_i^{(k+1)})), y_i^{(k+1)} - \tau\}$. It is easy to show that $\lambda_i \in [\lambda_i^-, \lambda_i^+]$ if τ and ε are small enough. With this modification we have reduced the cost of each iteration to just the computation of one of the two formulae in (5.3) and (5.4).

Iterative Scheme 3.

for $i = 1, \ldots, n$ **do**

$c_i^{(0)} = (x_i^{(0)} + y_i^{(0)})/2$

if $c_i^{(0)} > \lambda_i$ **then** $L(i) = 1, y_i^{(0)} = c_i^{(0)}$ **else** $L(i) = 0, x_i^{(0)} = c_i^{(0)}$

enddo

$S_1 = \{i : L(i) = 1\},\ S_2 = \{i : L(i) = 0\},\ k = 1$

while $S_1 \cup S_2 \neq 0$ **do**

for $i \in S_1$ **do**

$y_i^{(k+1)} = y_i^{(k)} - \frac{1}{P'(y_i^{(k)})/P(y_i^{(k)}) - \sum_{j=1,j\neq i}^{n} \frac{1}{y_i^{(k)} - x_j^{(k)}}}$, $x_i^{(k+1)} = x_i^{(k)}$

if $TEST(y_i^{(k+1)}, y_i^{(k)}) =' TRUE'$ **then** $S_1 = S_1 - \{i\}$

enddo

for $i \in S_2$ **do**

$x_i^{(k+1)} = x_i^{(k)} - \frac{1}{P'(x_i^{(k)})/P(x_i^{(k)}) - \sum_{j=1,j\neq i}^{n} \frac{1}{x_i^{(k)} - y_j^{(k)}}}$, $y_i^{(k+1)} = y_i^{(k)}$

if $TEST(x_i^{(k+1)}, x_i^{(k)}) =' TRUE'$ **then** $S_2 = S_2 - \{i\}$

enddo

for $i \notin S_1 \cup S_2$ **do**

$y_i^{(k+1)} = y_i^{(k)}, x_i^{(k+1)} = x_i^{(k)}$

enddo

$k = k + 1$

enddo

for $i \in \{1, \ldots, n\}$ **do**

if $L(i) = 1$

then $\lambda_i^+ = y_i^{(k)}, \lambda_i^- = \max\{y_i^{(k)}(1 - \varepsilon \,\mathrm{sign}(y_i^{(k)})), y_i^{(k)} - \tau\}$,

else $\lambda_i^- = x_i^{(k)}, \lambda_i^+ = \max\{x_i^{(k)}(1 + \varepsilon \,\mathrm{sign}(x_i^{(k)})), x_i^{(k)} + \tau\}$.

enddo

5.7 Numerical tests

Our algorithm is implemented and tested on SPARC stations with IEEE floating point standard. The machine precision is $\varepsilon \approx 2.2 \times 10^{-16}$.

Five types of matrices are used in testing our algorithms. They are

- The Toeplitz matrix $[1, 2, 1]$, i.e., all its diagonal elements are 2 and off-diagonal elements are 1;

- The random matrix with both diagonal and off-diagonal elements being uniformly distributed random numbers between 0 and 1;

- The Wilkinson matrix W_n^+, i.e., the matrix $[1, d_i, 1]$, where $d_i = |(n+1)/2 - i|$, $i = 1, 2, \ldots, n$ with n odd;
- The matrix $[1, \mu_i, 1]$, where $\mu_i = i \times 10^{-6}$;
- The matrix T_2: the same as matrix $[2, 8, 2]$ except the first diagonal element $\alpha_1 = 4$.

We compare the performance of the three algorithms: A01, A02 and A03 that use Iterative Scheme 1, Iterative Scheme2 and Iterative Scheme 3 respectively. The test results are listed in Table 5.1.

Matrix	Order n	A01	A02	A03
	63	0.04748	0.05665	0.04077
$[1, 2, 1]$	127	0.18626	0.21202	0.16487
	255	0.72778	0.85625	0.66848
	511	2.83899	3.38601	2.61953
	63	0.03550	0.03444	0.03357
Random	127	0.16321	0.16543	0.16020
	255	0.51073	0.51821	0.50249
	511	1.24389	1.19141	1.23842
	63	0.05840	0.05934	0.05244
W_n^+	127	0.19129	0.19708	0.17791
	255	0.64331	0.64409	0.59348
	511	1.98426	1.96220	1.86954
	63	0.07959	0.08513	0.06817
$[1, \mu_i, 1]$	127	0.32410	0.34346	0.27813
	255	1.29138	1.38410	1.15678
	511	5.21321	5.54776	4.78339
	63	0.07576	0.08429	0.06455
T_2	127	0.29829	0.33031	0.25644
	255	1.18816	1.31843	1.02157
	511	4.70662	5.17295	4.04890

Table 5.1: The execution time in seconds for evaluating all eigenvalues

Algorithm DETEVL:

`Input:` $T = [\beta_{i-1}, \alpha_i, \beta_i]$, λ

`Output:` $-\dfrac{P'(\lambda)}{P(\lambda)} = \eta_n$ **(where** $P(\lambda) \equiv \det(T - \lambda I)$**)**

and $\kappa(\lambda) = \mathit{neg\_count}$

(=the number of eigenvalues of T **less than** λ**)**

Begin DETEVL

$\xi_1 = \alpha_1 - \lambda$

If $\xi_1 = 0$ **then** $\xi_1 = \beta_1^2 \varepsilon^2$

$\eta_0 = 0, \quad \eta_1 = \dfrac{1}{\xi_1}, \quad \mathit{neg\_count} = 0$

If $\xi_1 < 0$ **then** $\mathit{neg\_count} = 1$

For $i = 2 : n$

$$\xi_i = (\alpha_i - \lambda) - \left(\frac{\beta_{i-1}^2}{\xi_{i-1}}\right)$$

If $\xi_i = 0$ **then** $\xi_i = \left(\dfrac{\beta_{i-1}^2}{\xi_{i-1}}\right)\varepsilon^2$

If $\xi_i < 0$ **then** $\mathit{neg\_count} = \mathit{neg\_count} + 1$

$$\eta_i = \frac{1}{\xi_i}\left[(\alpha_i - \lambda)\eta_{i-1} + 1 - \left(\frac{\beta_{i-1}^2}{\xi_{i-1}}\right)\eta_{i-2}\right]$$

End for

End DETEVL

Figure 5.1: **Algorithm** DETEVL

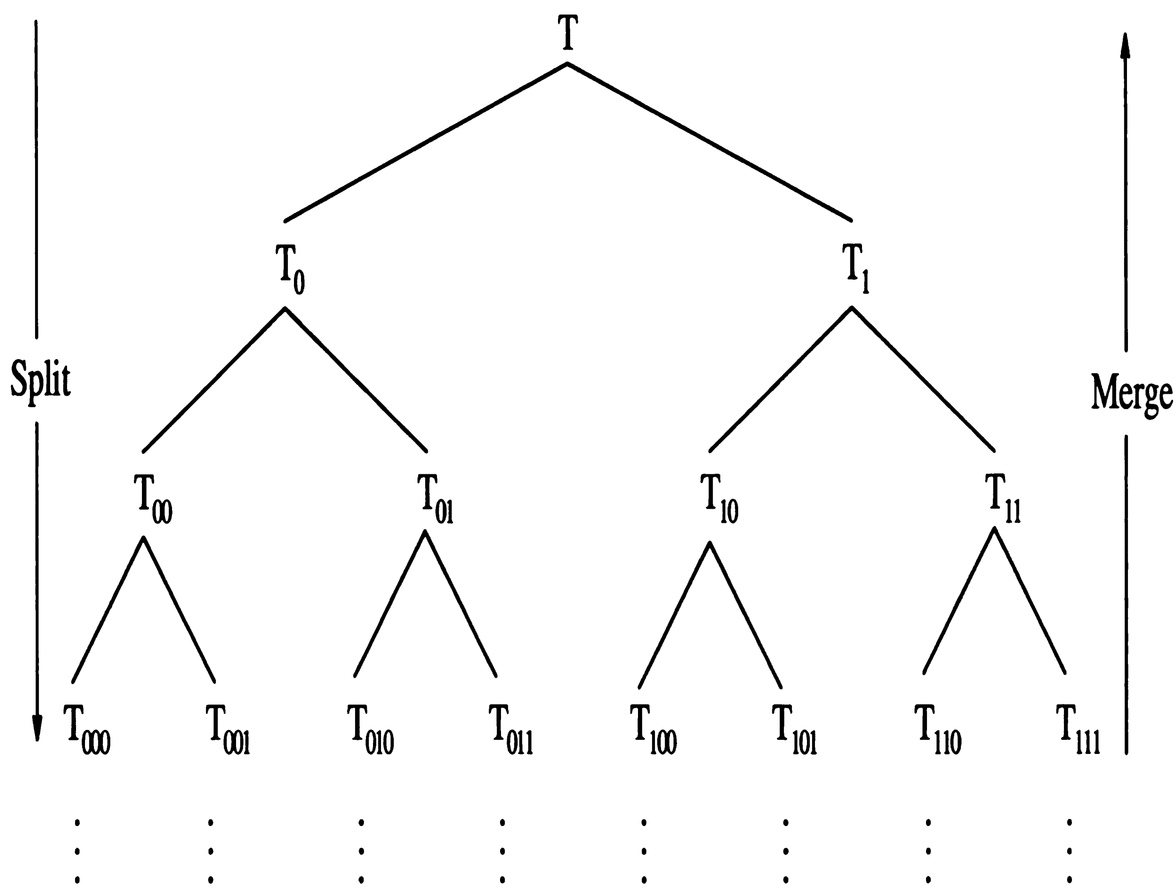

Figure 5.2: Split and merge processes

Bibliography

[1] Aberth, O. (1973), Iteration Methods for Finding all Zeros of a Polynomial Simultaneously, *Mathematics of Computation* **27**, 339–344.

[2] Alefeld, G. and J. Herzberger (1974), On the Convergence Speed of Some Algorithms for the Simultaneous Approximation of Polynomial Roots, *SIAM J. Numer.Anal.* **11**, 237–243.

[3] Alefeld, G. (1981), On the convergence of Halley's method, *Amer. Math. Monthly* **88**, 530–536.

[4] Anderson, E., Z. Bai, C. Bischof, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, S. Ostrouchov, and D. Sorenson (1992), *LAPACK User's Guide*, SIAM, Philadelphia.

[5] Bateman, H. (1938), Halley's methods for solving equations, *Amer. Math. Monthly* **45**, 11–17.

[6] Börsch-Supan, W. (1963), A posteriori error bounds for the zeros of polynomials, *Numer. Math.* **5**, 380–398.

[7] Braess, D. and K.P. Hadeler (1973), Simultaneous Inclusion of the Zeros of a Polynomial, *Numer. Math.* **21**, 161–165.

[8] Brown, G.H. , Jr. (1977), On Halley's variation of Newton's method, *Amer. Math. Monthly* **84**, 726–728.

[9] Cuppen, J.J.M. (1981), A divide and conquer method for the symmetric tridiagonal eigenproblem, *Numer. Math.* **36**, 177–195.

[10] Davies, M. and B. Dawson (1975), On the global convergence of Halley's iteration formula, *Numer. Math.* **24**, 133–135.

[11] Dieudonné, J. (1960), *Foundations of Modern Analysis*, Academic Press, New York.

[12] Dochev, K., (1962), An alternative method of Newton for simultaneous calculation of all the roots of a given algebraic equation, *Phys. Math. J. Bulgar. Acad. Sci.* **5**, 136–139.

[13] Dongarra, J.J. and D.C. Sorensen, (1987), A fully parallel algorithm for the symmetric eigenvalue problem, *SIAM J. Sci. Stat. Comput.* **8**, 139–154.

[14] Döring, B. (1970), Some theorems on the procedure of tangent hyperbolas (German), *Apl. Mat.* **15**, 418–464.

[15] Du, Q., M. Jin, T. Y. Li and Z. Zeng (1996), Quasi-Laguerre Iteration in Solving Symmetric Tridiagonal Eigenvalue Problems, *SIAM J. Sci. Comput.* **17**, No. 6, 1347–1368.

[16] Durand, I.E. (1960), Solutions Numérique des Équations Algébriques, Tome I: Équations du Type $F(x) = 0$; Racines d'une Polynôme, 279–281, Masson, Paris.

[17] Ehrlich, L.W. (1967), A modified Newton method for polynomials, *Comm. ACM* **10**, 107–108.

[18] Fraigniaud, P. (1991), The Durand-Kerner Polynomials Roots-Finding Method in Case of Multiple Roots, *BIT* **31**, 112–123.

[19] Frame, J.S. (1944), A variation of Newton's method, *Amer. Math. Monthly* **51**, 36–38.

[20] Frame, J.S. (1953), The solution of equations by continued fractions, *Amer. Math. Monthly* **60**, 293–305.

[21] Gander, W. (1985), On Halley's iteration method, *Amer. Math. Monthly* **92**, 131–134.

[22] Gargantini, I. and P. Henrici (1972), Circular arithmetic and the determination of polynomial zeros, *Numer. Math.* **18**, 305–320.

[23] Golub, G. H. and C.F. Van Loan (1989), *Matrix Computations, 2nd Ed.*, The John Hopkins University Press, Baltimore, MD.

[24] Gragg, W.B. and R.A. Tapia (1974), Optimal error bounds for the Newton-Kantorovich theorem, *SIAM J. Numer. Anal.* **11**, 10–13.

[25] Gregory, R.T. and D.L. Karney (1978), *A Collection of Matrices for Testing Computational Algorithms*, Robert E. Krieger Publishing Company, Huntington, New York.

[26] Gu, M. and S.C. Eisenstat (1992), A divide-and-conquer algorithm for the symmetric tridiagonal eigenproblem, *Research Report YALEU/DCS/RR-916*, Dept of Comput. Sci., Yale University.

[27] Hansen, E. and M. Patrick (1977), A family of root finding methods, *Numer. Math.* **27**, 257–269.

[28] Jankó, B. (1963), Sur une nouvelle généralisation de la méthode des hyperboles tangentes pour la résolution des équations functionnelles nonlinéaires définies dans l'espace de Banach, *Ann. Polon. Math.* **12**, 279–288.

[29] Kahan, W. (1967), Laguerre's method and a circle which contains at least one zero of a polynomial, *SIAM J. Numer. Anal.* **4** no. 3, 474–482.

[30] Kahan, W. (1992), Notes On Laguerre's Iteration, preprint, University of California, Berkeley.

[31] Kantorovich, L.V. and G. P. Akilov (1964), *Functional Analysis in Normed Spaces*, MacMillan, New York.

[32] Kerner, I.O. (1966), Ein Gesamtschrittverfahren zur Berechnung der Nullstellen von Polynomen, *Numerische Mathematik* **8** 290–294.

[33] Kiss, I. (1954), A generalization of Newton's approximation procedure (German), *Z. Angew. Math. Mech.* **34**, 68–69.

[34] Lang, S. (1983), *Real Analysis*, Addision-Wesley, Reading, Massachusetts.

[35] Li, K., T.Y. Li and Z. Zeng (1994), An Algorithm for the Generalized Symmetric Tridiagonal Eigenvalue Problem, preprint.

[36] Li, T.Y. and Z. Zeng (1994), Laguerre's iteration in solving the symmetric tridiagonal eigenproblem—revisited, *SIAM J. Sci. Comput.* **15** no. 5 , 1445–1473.

[37] Miel, G.J. (1979), The Kantorovich theorem with optimal error bounds, *Amer. Math. Monthly* **86**, 212–215.

[38] Ortega, J.M. and W.C. Rheinboldt (1970), *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York.

[39] Ostrowski, A.M. (1973), *Solution of Equations in Euclidean and Banach Spaces*, Academic Press, New York.

[40] Parlett, B. N.(1980), *The Symmetric Eigenvalue Problem*, Prentice-Hall, Englewood Cliffs, NJ.

[41] Petković, M. (1989), *Iterative Methods for Simultaneous Inclusion of Polynomial Zeros*, Lecture Notes in Mathematics, No.1387, Springer-Verlag, Berlin.

[42] Popovski, D.B. (1980), A family of one-point iteration formulae for finding roots, *Internat. J. Comput. Math.* **8**, 85–88.

[43] Rall, L.B. (1974), A Note On the Convergence of Newton's Method, *SIAM J. Numer. Anal.* **11** no. 1, 34–36.

[44] Šafiev, R.A. (1963a), The method of tangent hyperbolas, *Soviet Math. Dokl.* **4**, 482–485.

[45] Šafiev, R.A. (1963b), On a modification of the method of tangent hyperbolas (Russian), *Dokl. Akad. Nauk Azerbaidžan* **19**, 3–8.

[46] Scavo, T.R. and J.B. Thoo (1995), On the Geometry of Halley's Method, *Amer. Math. Monthly* **102** no. 5, 417–426.

[47] Schröder, E. (1870), On infinitely many algorithms for solving equations (German), *Mathematische Annalen* **2**, 317-365.

[48] Snyder, R.W. (1955), One more correction formula, *Amer. Math. Monthly* **62**, 722–725.

[49] Sorensen, D.C. and P.P.T. Tang (1991), On the orthogonality of eigenvectors computed by divide-and-conquer techniques, *SIAM. J. Numer. Anal.* **28**, 1752–1775.

[50] Stewart, J.K. (1951), Another variation of Newton's method, *Amer. Math. Monthly* **58**, 331–334.

[51] Traub, J.F. (1964), *Iterative Methods for the Solution of Equations*, Prentice-Hall, Englewwood Cliffs, NJ.

[52] Traub, J.F. and H. Wozniakowski (1979), Convergence and complexity of Newton iteration for operator equations, *J. Assoc. Comp. Mach.* **29**, 250–258.

[53] Trefftz, C., C.C. Huang, P. McKinley, T.Y. Li, and Z. Zeng (1995), A scalable eigenvalue solver for symmetric tridiagonal matrices, *Parallel Comput.* **21**, 1213–1240.

[54] Varga, R.S. (1962), *Matrix Iterative Analysis*, Prentice-Hall, Englewood Cliffs, N.J..

[55] Weierstrass, K. (1967), Newer Beweis des Satzes, dass jede ganze rationale Function einer Veränderlichen dargestellt werden kann als ein Product aus lineare Functionen derselben Veränderlichen, 251–269, Johnson Reprint Corp., New York.

[56] Wilkinson, J.H. (1965), *The Algebraic Eigenvalue Problem*, Oxford University Press, Oxford.

[57] Yamamoto, T., Kanno, S., Atanassova, L. (1994), Validated computation of polynomial zeros by the Durand-Kerner method, in: Herzberger, J., Ed., *Topics in Validated Computations*, North-Holland, Amsterdam, 27-53.

[58] Zheng, S. (1982), The convergence of Halley's method and its optimal error bounds (Chinese), *Journal of Hangzhou University* **9**, 285–289.