Journal of Computational and Applied Mathematics manuscript No.
(will be inserted by the editor)

Iteration Functions Re-visited

Michael Farmer, George Loizou, Stephen Maybank

Received: / Accepted:

Abstract Two classes of Iteration Functions (IFs) are derived in this paper. The first (one-point IFs) was originally derived by Joseph Traub using a different approach to ours (simultaneous IFs). The second is new and is demonstrably shown to be more *informationally efficient* than the first. These IFs apply to polynomials with arbitrary complex coefficients and zeros, which can also be multiple.

Keywords Iteration Functions · Polynomial Zeros · Multiple Zeros

Contents

1 Introduction

1.1 Context

It is better to start by putting the material in this paper into the context of our latest global algorithm [Far14, pp. 62–63]. We compute the zeros of arbitrary polynomials in two distinct stages.

Department of Computer Science and Information Systems
Birkbeck, University of London
WC1E 7HX, UK
E-mail: george@dcs.bbk.ac.uk
E-mail: sjmaybank@dcs.bbk.ac.uk

1.1.1 Stage 1

In this first stage we systematically search the complex plane for regions containing the zeros of a given polynomial. This search continues until we are satisfied that each region contains a single zero, which may be a multiple zero of the polynomial.

This search stage could be continued until we were satisfied that the centres of our regions were sufficiently accurate approximations to the true values of the zeros. However, the work is computationally intensive, so we switch to Stage two, which offers quicker convergence.

1.1.2 Stage 2

This second stage uses the centres of our regions found in Stage 1 as initial approximations for IFs that converge rapidly to accurate approximations to the true values of the zeros of a polynomial defined over the complex numbers or the real numbers.

It is the derivation of and discussion about these IFs that form the basis of this paper. The IFs are used for computing the zeros of arbitrary polynomials given suitable initial values for which convergence can be achieved. The main justification for this paper is to present our contention that working with multiple zeros is the best approach [FL85]; simple zeros are just a special case.

The second *reason d'être* is to present some results that were missing from the first named author's recent PhD thesis [Far13], namely the exact Asymptotic Error Constants (AECs) of the fourth order and fifth order simultaneous IFs presented therein. In addition, the AEC of the third order IF in (A.35) of [Far13] is corrected.

The third reason is to present our approach using IFs in *polynomial format*, which emphasises more clearly how the structure of the IFs changes as we build IFs of higher and higher orders, yielding a power series obtained from our given polynomial.

The results of this paper have been incorporated into a revised version of the thesis. This version is available on the web [Far14]. All references to the thesis are to the web version rather than the original submission.

The remainder of this paper is set out as follows.

1.2 The Scheme of Things

Next, §2 takes the reader through definitions and equations that are used throughout this paper.

This is followed by §3 which derives a class of one-point IFs followed by a class of simultaneous IFs.

Next, §4 derives the orders of convergence of the various IFs, and especially their AECs. The IFs for polynomials with only simple zeros are easily derived from the IFs for polynomials with multiple zeros.

Next, §5 contains some remarks concerning R-order convergence, where applying an IF in a *serial* fashion, rather than in a *parallel* fashion, can improve the order of convergence in certain cases.

An overview of Stage 1 of the algorithm for finding the zeros of a polynomial is given in §6 and information about the software and the experiments is given in §7.

Finally, §8 presents a summary of our conclusions on the work presented in this paper.

2 Preamble

Let $p(z)$ be a polynomial of degree n given by

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0, \;\; a_n a_0 \neq 0 \;\; , \tag{1}$$

where the coefficients have complex or real values and the zeros also have complex or real values. In addition, the zeros $\{\alpha_i\}$ can have multiplicities $\{m_i\}$ greater than one, respectively. The polynomial $p(z)$ has N distinct zeros.

Throughout this paper sums and products will be over the range $1,\ 2,\ \ldots,\ N$ ($N \in \mathbb{N}^+$, the set of positive integers) unless stated otherwise, e.g.

$$\begin{aligned} \sum_{i<\nu} a_i &\equiv \sum_{i=1}^{\nu-1} a_i \;\; , \\ \prod_{i\neq\nu} b_i &\equiv \prod_{\substack{i=1 \\ i\neq\nu}}^{N} b_i \;\; . \end{aligned} \tag{2}$$

When $p(z)$ is **monic**, $a_n = 1$, and we have

$$p(z) = \prod_i (z-\alpha_i)^{m_i} \;\; . \tag{3}$$

We next define

$$M = \max_i m_i \;\; . \tag{4}$$

Let $p(z)$ be defined as in Equation (1). The following definitions, originally given by Joseph Traub [Tra82, pp. 5–6], will be used subsequently.

$$u(z) = \frac{p(z)}{p'(z)} \;\; , \tag{5}$$

which Joseph Traub calls the *normalised* $p(z)$, and

$$A_i(z) = \frac{p^{(i)}(z)}{i!\,p'(z)},\; i = 1,\ 2,\ \ldots,\ n \;\; , \tag{6}$$

where $p^{(i)}(z)$ is the ith derivative of $p(z)$ and which Joseph Traub calls the *normalised Taylor series coefficient*. Note that $u(z)$ is often referred to as **Newton's correction** [Pet89, p. 85]. For later use it is worth noting that

$$A_i'(z) = (i+1)A_{i+1}(z) - 2A_2(z)A_i(z),\; i = 1,\ 2,\ \ldots,\ n \;\; . \tag{7}$$

The following definitions will also be useful.

$$S_k(z_\nu) = \sum_{i \neq \nu} \frac{m_i}{(z_\nu - \alpha_i)^k},\ k = 1,\ 2,\ \ldots,\ N\ . \tag{8}$$

We note that

$$S_k'(z_\nu) = -kS_{k+1}(z_\nu),\ k = 1,\ 2,\ \ldots,\ N\ . \tag{9}$$

Let z_i be an approximation to the zero α_i, and $\hat{z}_i$ be the next approximation to α_i, using some iterative scheme. We now define

$$\begin{aligned} \varepsilon_i &= z_i - \alpha_i, \\ \hat{\varepsilon}_i &= \hat{z}_i - \alpha_i, \end{aligned} \quad i = 1,\ 2,\ \ldots,\ N\ , \tag{10}$$

with

$$\varepsilon = \max_i |\varepsilon_i|,\ \ i = 1,\ 2,\ \ldots,\ N\ . \tag{11}$$

3 Multiple Zeros

This section describes our two classes of IFs. We begin with some useful basic equations.

3.1 Basic Equations

In [FL75] we derived a class of IFs for improving the zeros of a polynomial with only simple zeros. In a later paper [FL77] we extended those results to multiple zeros in the context of a *globally convergent* algorithm [Far14]. This section presents a new class of IFs expressed in a *polynomial format* in preference to the *rational format* used previously by us and other authors, see [FL77], [Kis54], and [Tra82] for examples.

If $p(z)$, as defined in Equation (3), has a zero α_ν with multiplicity m_ν then the function $P(z_\nu)$ defined by

$$P(z_\nu) = p^{\frac{1}{m_\nu}}(z_\nu) \tag{12}$$

has a simple zero, α_ν. The following definitions will be used subsequently.

$$U(z_\nu) = \frac{P(z_\nu)}{P'(z_\nu)} = m_\nu u(z_\nu)\ \ , \tag{13}$$

$$B_i(z_\nu) = \frac{P^{(i)}(z_\nu)}{i!P'(z_\nu)}\ \ i = 1,\ 2,\ \ldots,\ N. \tag{14}$$

Using Equations (12) and (14), it is simple to verify that

$$B_i'(z_\nu) = (i+1)B_{i+1}(z_\nu) - 2B_2(z_\nu)B_i(z_\nu)\ \ i = 1,\ 2,\ \ldots,\ N, \tag{15}$$

which will be used subsequently. Once again, following our approach in [FL75], we can use the Taylor series expansion of $P(\alpha_\nu)$ about z_ν [Hen74, pp. 492–493], since $P(z_\nu)$ has a simple zero α_ν, to obtain, after dividing both sides of the series by $P'(z_\nu)$,

$$U(z_\nu) = \sum_{i=1}^{\rho-1} (-1)^{i-1} B_i(z_\nu)\varepsilon_\nu^i + O(\varepsilon_\nu^\rho),\ (\rho \in \mathbb{N}^+) \ . \tag{16}$$

Equation (16) can be rearranged to obtain a class of IFs. We start with $\rho = 2$ (see Equations (19) and (21)), namely

$$\Psi_2(z_\nu) = U(z_\nu) \ . \tag{17}$$

Now, as ρ increases, substitute ε_ν^i in Equation (16) with the appropriate value for powers of $\Psi_i(z_\nu)$ as follows

$$\begin{aligned}
\Psi_2(z_\nu) &= U(z_\nu) \ , \\
\Psi_3(z_\nu) &= U(z_\nu) + B_2(z_\nu)\Psi_2^2(z_\nu) \ , \\
\Psi_4(z_\nu) &= U(z_\nu) + B_2(z_\nu)\Psi_3^2(z_\nu) - B_3(z_\nu)\Psi_2^3(z_\nu) \ , \\
\Psi_5(z_\nu) &= U(z_\nu) + B_2(z_\nu)\Psi_4^2(z_\nu) - B_3(z_\nu)\Psi_3^3(z_\nu) + B_4(z_\nu)\Psi_2^4(z_\nu) \ ,
\end{aligned} \tag{18}$$

etc. Then the IF

$$R_\rho = z_\nu - \Psi_\rho(z_\nu) \tag{19}$$

is of order ρ. If the higher-order powers of $U(z_\nu)$ other than those required for the IFs (of the appropriate order) are removed, then the expansion of Equation (18) yields a class of IFs which we term **basic**, as they form the basis for our new class.

$$\begin{aligned}
\Lambda_2(z_\nu) &= U(z_\nu) \ , \\
\Lambda_3(z_\nu) &= U(z_\nu) + B_2(z_\nu)U^2(z_\nu) \ , \\
\Lambda_4(z_\nu) &= U(z_\nu) + B_2(z_\nu)U^2(z_\nu) + [2B_2^2(z_\nu) - B_3(z_\nu)]U^3(z_\nu) \ , \\
\Lambda_5(z_\nu) &= U(z_\nu) + B_2(z_\nu)U^2(z_\nu) + [2B_2^2(z_\nu) - B_3(z_\nu)]U^3(z_\nu) \\
&\quad + [5B_2^3(z_\nu) - 5B_2(z_\nu)B_3(z_\nu) + B_4(z_\nu)]U^4(z_\nu) \ .
\end{aligned} \tag{20}$$

The corresponding IF is

$$R_\rho = z_\nu - \Lambda_\rho(z_\nu). \tag{21}$$

As mentioned in [FL77, pp. 428–429], some members of this class of IFs are well known when only simple zeros are present. They are also described in [Far14, pp. 47–53]. The IFs (19) and (21) are identical to within an error of order $O(\varepsilon^\rho)$, in that

$$\Lambda_\rho(z_\nu) = \Psi_\rho(z_\nu) + O(\varepsilon_\nu^\rho),$$

however they have different asymptotic error constants.

The IFs presented below are given in *polynomial format* rather than the more common *rational format*. We think that this emphasises how the IFs in the different classes are generated as their order of convergence increases. To see the rational format IF corresponding to one of our current IFs, see [FL77, pp.429–430], [Far14, pp.122–123], and [Kis54, p. 68].

3.2 One-point Iteration Functions

We consider it necessary to include the older one-point IFs as they form an intermediary stage between the **basic** equations and our newly derived simultaneous IFs.

From a computational point of view $B_i(z_\nu)$ cannot be calculated directly from $p(z_\nu)$. Therefore we require an equation converting $B_i(z_\nu)$ to $A_i(z_\nu)$.

In order to express the basic equations given in Equation (20) in terms of $u(z_\nu)$ and $A_i(z_\nu)$, we successively differentiate Equation (12) to obtain the following set of equations.

$$\begin{aligned}
B_2(z_\nu) &= -\frac{m_\nu - 1}{2!m_\nu u(z_\nu)} + A_2(z_\nu) \ , \\
B_3(z_\nu) &= \frac{(m_\nu - 1)(2m_\nu - 1)}{3!m_\nu^2 u^2(z_\nu)} - \frac{m_\nu - 1}{m_\nu u(z_\nu)} A_2(z_\nu) + A_3(z_\nu) \ , \\
B_4(z_\nu) &= -\frac{(m_\nu - 1)(2m_\nu - 1)(3m_\nu - 1)}{4!m_\nu^3 u^3(z_\nu)} \\
&\quad + \frac{(m_\nu - 1)(2m_\nu - 1)}{2!m_\nu^2 u^2(z_\nu)} A_2(z_\nu) \\
&\quad - \frac{(m_\nu - 1)}{m_\nu u(z_\nu)} \left[\frac{A_2^2(z_\nu)}{2!} + A_3(z_\nu) \right] + A_4(z_\nu) \ ,
\end{aligned} \tag{22}$$

etc. which gives us another class of IFs depending explicitly on m_ν, the multiplicity of the zero α_ν. Note that these equations have been verified using a Matlab program, `multiple.m`, see §7. These are described below.

3.2.1 Ernst Schröder's second-order modified Newton IF [Sch70]

$$\hat{z}_\nu = z_\nu - m_\nu u(z_\nu) \ . \tag{23}$$

3.2.2 Joseph Traub's third-order IF [Tra82, p. 139]

$$\hat{z}_\nu = z_\nu + m_\nu \left(\frac{m_\nu - 3}{2} \right) u(z_\nu) - m_\nu^2 A_2(z_\nu) u^2(z_\nu) \ , \tag{24}$$

which might be better known in its *rational* format, as Ljiljana Petković, Miodrag Petković, and Dragan Živković have demonstrated that this family is a form of Laguerre's method [PPŽ03, pp. 111–112]. For further information about Laguerre's method, especially in the case of real zeros, we recommend Alston Householder's description in [Hou70, pp.176–179].

3.2.3 Joseph Traub's fourth-order IF [Tra82, p. 139]

$$
\begin{aligned}
\hat{z}_\nu &= z_\nu - m_\nu \left(\frac{m_\nu^2 - 6m_\nu + 11}{6} \right) u(z_\nu) + m_\nu^2 (m_\nu - 2) A_2(z_\nu) u^2(z_\nu) \\
&\quad - m_\nu^3 [2A_2^2(z_\nu) - A_3(z_\nu)] u^3(z_\nu) \ , \qquad (25)
\end{aligned}
$$

which he originally defined in a variation of a polynomial format referred to as the **Horner** format after the well-known Horner's rule for the efficient evaluation of a polynomial [Hou70, pp. 3–4].

3.2.4 Joseph Traub's fifth-order IF [Tra82, p. 139]

$$
\begin{aligned}
\hat{z}_\nu &= z_\nu + 24^{-1} m_\nu (m_\nu^3 - 10m_\nu^2 + 35m_\nu - 50) u(z_\nu) \\
&\quad - 12^{-1} m_\nu^2 (7m_\nu^2 - 30m_\nu + 35) A_2(z_\nu) u^2(z_\nu) \\
&\quad + 2^{-1} m_\nu^3 (3m_\nu - 5) [2A_2^2(z_\nu) - A_3(z_\nu)] u^3(z_\nu) \\
&\quad - m_\nu^4 [5A_2^3(z_\nu) - 5A_2(z_\nu)A_3(z_\nu) + A_4(z_\nu)] u^4(z_\nu) \ , \qquad (26)
\end{aligned}
$$

which he originally defined in Horner format.

3.3 Simultaneous Iteration Functions

To improve the informational efficiency of our one-point IFs, we replace the highest-order derivative in each IF with an expression containing only lower-order derivatives. We start as follows.

$$
\begin{aligned}
\frac{1}{u(z_\nu)} &= \frac{p'(z_\nu)}{p(z_\nu)} \ , \\
&= \frac{\sum_i m_i (z_\nu - \alpha_i)^{m_i - 1} \prod_{j \neq i} (z_\nu - \alpha_j)^{m_j}}{\prod_j (z_\nu - \alpha_j)^{m_j}} \ , \\
&= \sum_i \frac{m_i}{(z_\nu - \alpha_i)} \ , \\
&= \frac{m_\nu}{\varepsilon_\nu} + S_1(z_\nu) \ . \qquad (27)
\end{aligned}
$$

We also have, by a Taylor series expansion,

$$
\begin{aligned}
\frac{1}{U(z_\nu)} &= \frac{1}{\varepsilon_\nu} \{ 1 + B_2(z_\nu)\varepsilon_\nu + [B_2^2(z_\nu) - B_3(z_\nu)]\varepsilon_\nu^2 \\
&\quad + [B_2^3(z_\nu) - 2B_2(z_\nu)B_3(z_\nu) + B_4(z_\nu)]\varepsilon_\nu^3 \\
&\quad + [B_2^4(z_\nu) - 3B_2^2(z_\nu)B_3(z_\nu) + 2B_2(z_\nu)B_4(z_\nu) + B_3^2(z_\nu) - B_5(z_\nu)]\varepsilon_\nu^4 \} \\
&\quad + O(\varepsilon_\nu^4) \ . \qquad (28)
\end{aligned}
$$

Combining Equations (13), (27) and (28) we obtain

$$\begin{aligned} S_1(z_\nu) = m_\nu\{&B_2(z_\nu)+[B_2^2(z_\nu)-B_3(z_\nu)]\varepsilon_\nu+[B_2^3(z_\nu)-2B_2(z_\nu)B_3(z_\nu)+B_4(z_\nu)]\varepsilon_\nu^2 \\ &+[B_2^4(z_\nu)-3B_2^2(z_\nu)B_3(z_\nu)+2B_2(z_\nu)B_4(z_\nu)+B_3^2(z_\nu)-B_5(z_\nu)]\varepsilon_\nu^3\} \\ &+O(\varepsilon_\nu^4) \ . \end{aligned} \tag{29}$$

Differentiating both sides of Equation (29) with respect to z_ν we now obtain

$$\begin{aligned} S_1'(z_\nu) = m_\nu\{&2B_3(z_\nu)-B_2^2(z_\nu)-2[B_2^3(z_\nu)-2B_2(z_\nu)B_3(z_\nu)+B_4(z_\nu)]\varepsilon_\nu \\ &-[3B_2^4(z_\nu)-8B_2^2(z_\nu)B_3(z_\nu)+3B_3^2(z_\nu)+4B_2(z_\nu)B_4(z_\nu)-2B_5(z_\nu)]\varepsilon_\nu^2\} \\ &+O(\varepsilon_\nu^3) \ . \end{aligned} \tag{30}$$

Continuing in the same vein we obtain

$$\begin{aligned} S_1''(z_\nu) = m_\nu\{&2B_2^3(z_\nu)-6B_2(z_\nu)B_3(z_\nu)+6B_4(z_\nu) \\ &+6[B_2^4(z_\nu)-3B_2^2(z_\nu)B_3(z_\nu)+B_3^2(z_\nu)+2B_2(z_\nu)B_4(z_\nu)-B_5(z\nu)]\varepsilon_\nu\} \\ &+O(\varepsilon_\nu^2) \ . \end{aligned} \tag{31}$$

Finally, combining Equations (29) through (31) together with Equation (9) yields the following

$$\begin{aligned} S_1(z_\nu) &= m_\nu B_2(z_\nu)+O(\varepsilon) \ , \\ S_2(z_\nu) &= m_\nu[B_2^2(z_\nu)-2B_3(z_\nu)]+O(\varepsilon) \ , \\ S_3(z_\nu) &= m_\nu[B_2^3(z_\nu)-3B_2(z_\nu)B_3(z_\nu)+3B_4(z_\nu)]+O(\varepsilon) \ . \end{aligned} \tag{32}$$

In our basic IF of order ρ, as given by (21), we cannot replace $B_{\rho-1}(z_\nu)$ with one of the above equations containing $S_{\rho-2}(z_\nu)$ because every $S_k(z_\nu)$ contains terms involving the unknown zeros, α_i. We overcome this problem by introducing the following definition.

$$T_k(z_\nu) = \sum_{i\neq\nu}\frac{m_i}{(z_\nu-z_i)^k}. \tag{33}$$

Replacing z_i by $\alpha_i+\varepsilon_i$, from Equation (10), yields the following.

$$\begin{aligned} T_k(z_\nu) &= \sum_{i\neq\nu}\frac{m_i}{(z_\nu-\alpha_i-\varepsilon_i)^k} \ , \\ &= \sum_{i\neq\nu}\frac{m_i}{(z_\nu-\alpha_i)^k(1-\frac{\varepsilon_i}{z_\nu-\alpha_i})^k} \ , \\ &= \sum_{i\neq\nu}\frac{m_i}{(z_\nu-\alpha_i)^k}(1-\frac{\varepsilon_i}{z_\nu-\alpha_i})^{-k} \ , \\ &= S_k(z_\nu)+O(\varepsilon) \ , \end{aligned} \tag{34}$$

which gives the required equations, namely

$$\begin{aligned} T_1(z_\nu) &= m_\nu B_2(z_\nu)+O(\varepsilon) \ , \\ T_2(z_\nu) &= m_\nu[B_2^2(z_\nu)-2B_3(z_\nu)]+O(\varepsilon) \ , \\ T_3(z_\nu) &= m_\nu[B_2^3(z_\nu)-3B_2(z_\nu)B_3(z_\nu)+3B_4(z_\nu)]+O(\varepsilon) \ , \end{aligned} \tag{35}$$

on using Equations (32). These equations are therefore used to replace the highest-order derivative in each of the basic IFs, given in Equation (20), thus generating the new IFs

$$R_\rho = z_v - \Xi_\rho(z_v),$$

in which the terms $\Xi_\rho(z_v)$ are given by

$$\begin{aligned}
\Xi_2(z_v) &= U(z_v) \ , \\
\Xi_3(z_v) &= U(z_v) + m_v^{-1} T_1(z_v) U^2(z_v) \ , \\
\Xi_4(z_v) &= U(z_v) + B_2(z_v) U^2(z_v) + [(3/2) B_2^2(z_v) + (2m_v)^{-1} T_2(z_v)] U^3(z_v)) \ , \\
\Xi_5(z_v) &= U(z_v) + B_2(z_v) U^2(z_v) + [2B_2^2(z_v) - B_3(z_v)] U^3(z_v) \\
&\quad + [(14/3) B_2^3(z_v) - 4B_2(z_v) B_3(z_v) + (3m_v)^{-1} T_3(z_v)] U^4(z_v) \ . \qquad (36)
\end{aligned}$$

Once again, note that these new IFs have been verified by a Mathematica version 10 program, see §7. There is no second-order simultaneous IF expressible in *polynomial format*.

3.3.1 *Our third-order IF*

$$\hat{z}_v = z_v - m_v u(z_v) - m_v T_1(z_v) u^2(z_v) \ . \qquad (37)$$

3.3.2 *Our fourth-order IF*

$$\begin{aligned}
\hat{z}_v &= z_v - \frac{m_v}{8}(3m_v^2 - 10m_v + 15) u(z_v) + \frac{m_v^2}{2}(3m_v - 5) A_2(z_v) u^2(z_v) \\
&\quad - \frac{m_v^2}{2}[3m_v A_2^2(z_v) + T_2(z_v)] u^3(z_v) \ . \qquad (38)
\end{aligned}$$

3.3.3 *Our fifth-order IF*

$$\begin{aligned}
\hat{z}_v = z_v &- \frac{m_v}{12}(m_v^3 + 3m_v^2 - 17m_v + 25) u(z_v) \\
&- \frac{m_v^2}{6}(m_v^2 - 12m_v + 17) A_2(z_v) u^2(z_v) \\
&+ m_v^3[(3m_v - 5) A_2^2(z_v) - (2m_v - 3) A_3(z_v)] u^3(z_v) \\
&- \frac{m_v^3}{3}[14 m_v A_2^3(z_v) - 12 m_v A_2(z_v) A_3(z_v) + T_3(z_v)] u^4(z_v) \ .
\end{aligned} \qquad (39)$$

3.4 Efficiency Considerations

According to Joseph Traub [Tra82, pp. 11–13] the *informational efficiency*, EFF, of an IF is the order of the IF, i.e. ρ in our notation, divided by the *informational usage* of the IF, i.e. d in our case, namely the number of new polynomial and polynomial derivative evaluations required per iteration. Thus

$$EFF = \frac{\rho}{d} \ . \tag{40}$$

In addition, he also proves that $EFF \leq 1$ for one-point IFs. This leads him to define a one-point IF as *optimal* if its $EFF = 1$. It follows therefore that the one-point IFs, described in §3.2, are optimal.

However, our simultaneous IFs, as described in §3.3, have an EFF of

$$EFF = \frac{\rho}{d-1} \tag{41}$$

because the highest-order derivative in each basic IF has been replaced by an equation using only lower-order derivatives. Therefore, our simultaneous IFs, described in §3.3, are more informationally efficient than the one-point IFs, described in §3.2. These results are summarised in Table 3.4 below.

Order	**One-point IFs**	**Simultaneous IFs**
2	1	n/a
3	1	1.5
4	1	1.33
5	1	1.25

Table 1. Comparative Efficiency of IFs

The paper by Robert Voigt [Voi71] contains some interesting insights into this topic.

4 Convergence of IFs

There is a certain symmetry between those equations used to generate our IFs and those used to generate our asymptotic error constants (AECs). When generating our IFs we have

$$p(\alpha_\nu) = p(z_\nu) - p'(z_\nu)\varepsilon_\nu + \frac{p''(z_\nu)}{2!}\varepsilon_\nu^2 - \cdots \ . \tag{42}$$

Since $p(\alpha_\nu) = 0$ we have

$$p(z_\nu) = p'(z_\nu)\varepsilon_\nu - \frac{p''(z_\nu)}{2!}\varepsilon_\nu^2 + \cdots \ . \tag{43}$$

However, when generating our AECs we have

$$\begin{aligned}\tilde{p}(z_\nu) &= p(\alpha_\nu) + p'(\alpha_\nu)\varepsilon_\nu + \frac{p''(\alpha_\nu)}{2!}\varepsilon_\nu^2 + \cdots \ , \\ &= p'(\alpha_\nu)\varepsilon_\nu + \frac{p''(\alpha_\nu)}{2!}\varepsilon_\nu^2 + \cdots \ .\end{aligned} \tag{44}$$

In order to distinguish between the two expansions of $p(z_\nu)$ we designate the expansion given in Equation (44) as $\tilde{p}(z_\nu)$. This notation will also be used for the other functions used in generating our AECs, see §4.1.

4.1 Basic Equations

The following definition is a slight variation on Joseph Traub's $B_{j,m}(z)$ which he calls the *generalised normalised Taylor series coefficient* [Tra82, p. 6],

$$C_i(z_\nu) = \frac{1}{m_\nu}\frac{p^{(m_\nu+i-1)}(z_\nu)}{(m_\nu+i-1)!}\frac{(m_\nu)!}{p^{(m_\nu)}(z_\nu)} \ , \ i = 1,\ 2,\ \ldots \ , \tag{45}$$

which is used when the multiplicity $m_\nu > 1$. Note that when only simple zeros are present, then

$$A_i(z_\nu) \equiv B_i(z_\nu) \equiv C_i(z_\nu), \ \ i = 1,\ 2,\ \ldots,\ n \ . \tag{46}$$

In the case of a multiple zero, α_ν, we have

$$\tilde{p}(z_\nu) = \sum_{i=m_\nu}^{\infty} \frac{p^{(i)}(\alpha_\nu)}{i!}\varepsilon_\nu^i \ , \tag{47}$$

Now multiply the above by the factor

$$\kappa(\varepsilon_\nu) = \frac{m_\nu!}{m_\nu p^{(m_\nu)}(\alpha_\nu)\varepsilon_\nu^{m_\nu-1}} \tag{48}$$

to obtain

$$\kappa(\varepsilon_\nu)\tilde{p}(z_\nu) = \frac{\varepsilon_\nu}{m_\nu} + C_2(\alpha_\nu)\varepsilon_\nu^2 + C_3(\alpha_\nu)\varepsilon_\nu^3 + C_4(\alpha_\nu)\varepsilon_\nu^4 + C_5(\alpha_\nu)\varepsilon_\nu^5 + \ldots \tag{49}$$

The derivative of $\tilde{p}(z_\nu)$ is obtained using the equation $\varepsilon_\nu = z_\nu - \alpha_\nu$ and the identity

$$\kappa(\varepsilon_\nu)\tilde{p}'(z_\nu) = (\kappa(\varepsilon_\nu)\tilde{p}(z_\nu))' - \kappa'(\varepsilon_\nu)\tilde{p}(z_\nu) \ .$$

The higher order derivatives of $\tilde{p}(z_\nu)$ are obtained in a similar way.

The factor $\kappa(\varepsilon_\nu)$ cancels out in applications. For example,

$$\tilde{u}(z_\nu) = \frac{\tilde{p}(z_\nu)}{\tilde{p}'(z_\nu)} =$$

$$\varepsilon_\nu m_\nu^{-1}\{1 - C_2(\alpha_\nu)\varepsilon_\nu + [(m_\nu+1)C_2^2(\alpha_\nu) - 2C_3(\alpha_\nu)]\varepsilon_\nu^2
- [(m_\nu^2+2m_\nu+1)C_2^3(\alpha_\nu) - (3m_\nu+4)C_2(\alpha_\nu)C_3(\alpha_\nu) + 3C_4(\alpha_\nu)]\varepsilon_\nu^3
+ [(m_\nu^3+3m_\nu^2+3m_\nu+1)C_2^4(\alpha_\nu) - 2(2m_\nu^2+5m_\nu+3)C_2^2(\alpha_\nu)C_3(\alpha_\nu)]\varepsilon_\nu^4
+ [2(2m_\nu+3)C_2(\alpha_\nu)C_4(\alpha_\nu) + 2(m_\nu+2)C_3^2(\alpha_\nu) - 4C_5(\alpha_\nu)]\varepsilon_\nu^4\}
+ O(\varepsilon_\nu^6) \ . \tag{50}$$

Continuing in a similar vein, we have

$$\tilde{A}_2(z_\nu) = \frac{\tilde{p}''(z_\nu)}{2!\tilde{p}'(z_\nu)} \ ,
\tilde{A}_3(z_\nu) = \frac{\tilde{p}'''(z_\nu)}{3!\tilde{p}'(z_\nu)} \ ,
\tilde{A}_4(z_\nu) = \frac{\tilde{p}''''(z_\nu)}{4!\tilde{p}'(z_\nu)} \ , \tag{51}$$

etc. Evaluating the right-hand sides of Equation (51) finally yields the following equations.

$$\tilde{A}_2(z_\nu) = (2!\varepsilon_\nu)^{-1}\{[m_\nu-1] + [m_\nu+1]C_2(\alpha_\nu)\varepsilon_\nu
- [(m_\nu^2+2m_\nu+1)C_2^2(\alpha_\nu) - 2(m_\nu+2)C_3(\alpha_\nu)]\varepsilon_\nu^2
+ [(m_\nu^3+3m_\nu^2+3m_\nu+1)C_2^3(\alpha_\nu) - 3(m_\nu^2+3m_\nu+2)C_2(\alpha_\nu)C_3(\alpha_\nu)
+ 3(m_\nu+3)C_4(\alpha_\nu)]\varepsilon_\nu^3
- [(m_\nu^4+4m_\nu^3+6m_\nu^2+4m_\nu+1)C_2^4(\alpha_\nu)
- 4(m_\nu^3+4m_\nu^2+5m_\nu+2)C_2^2(\alpha_\nu)C_3(\alpha_\nu) + 4(m_\nu^2+4m_\nu+3)C_2(\alpha_\nu)C_4(\alpha_\nu)
+ 2(m_\nu^2+4m_\nu+4)C_3^2(\alpha_\nu) - 4(m_\nu+4)C_5(\alpha_\nu)]\varepsilon_\nu^4\} + O(\varepsilon_\nu^4) \ , \tag{52}$$

$$\tilde{A}_3(z_\nu) = (3!\varepsilon_\nu^2)^{-1}\{[m_\nu^2-3m_\nu+2] + 2(m_\nu^2-1)C_2(\alpha_\nu)\varepsilon_\nu
- [2(m_\nu^3+m_\nu^2-m_\nu-1)C_2^2(\alpha_\nu) - 2(2m_\nu^2+3m_\nu-2)C_3(\alpha_\nu)]\varepsilon_\nu^2
+ [2(m_\nu^4+2m_\nu^3-2m_\nu-1)C_2^3(\alpha_\nu) - 2(3m_\nu^3+7m_\nu^2-4)C_2(\alpha_\nu)C_3(\alpha_\nu)
+ 6(m_\nu^2+3m_\nu)C_4(\alpha_\nu)]\varepsilon_\nu^3
- [2(m_\nu^5+3m_\nu^4+2m_\nu^3-2m_\nu^2-3m_\nu-1)C_2^4(\alpha_\nu)
- 2(4m_\nu^4+13m_\nu^3+8m_\nu^2-7m_\nu-6)C_2^2(\alpha_\nu)C_3(\alpha_\nu)
+ 2(4m_\nu^3+15m_\nu^2+8m_\nu-3)C_2(\alpha_\nu)C_4(\alpha_\nu)
+ 2(2m_\nu^3+7m_\nu^2+4m_\nu-4)C_3^2(\alpha_\nu) - 4(2m_\nu^2+9m_\nu+4)C_5(\alpha_\nu)]\varepsilon_\nu^4\}
+ O(\varepsilon_\nu^3) \ , \tag{53}$$

$$
\begin{aligned}
\tilde{A}_4(z_\nu) = {} & (4!\varepsilon_\nu^3)^{-1}\{[m_\nu^3-6m_\nu^2+11m_\nu-6]+[(3m_\nu^3-6m_\nu^2-3m_\nu+6)C_2(\alpha_\nu)]\varepsilon_\nu \\
& -[(3m_\nu^4-3m_\nu^3-9m_\nu^2+3m_\nu+6)C_2^2(\alpha_\nu)-(6m_\nu^3-18m_\nu+12)C_3(\alpha_\nu)]\varepsilon_\nu^2 \\
& +[3(m_\nu^5-4m_\nu^3-2m_\nu^2+3m_\nu+2)C_2^3(\alpha_\nu) \\
& -3(3m_\nu^4+2m_\nu^3-11m_\nu^2-2m_\nu+8)C_2(\alpha_\nu)C_3(\alpha_\nu) \\
& +3(3m_\nu^3+6m_\nu^2-7m_\nu+6)C_4(\alpha_\nu)]\varepsilon_\nu^3 \\
& -[3(m_\nu^6+m_\nu^5-4m_\nu^4-6m_\nu^3+m_\nu^2+5m_\nu+2)C_2^4(\alpha_\nu) \\
& -6(2m_\nu^5+3m_\nu^4-7m_\nu^3-9m_\nu^2+5m_\nu+6)C_2^2(\alpha_\nu)C_3(\alpha_\nu) \\
& +6(2m_\nu^4+5m_\nu^3-4m_\nu^2-m_\nu+6)C_2(\alpha_\nu)C_4(\alpha_\nu) \\
& +6(m_\nu^4+2m_\nu^3-3m_\nu^2-4m_\nu+4)C_3^2(\alpha_\nu) \\
& -12(m_\nu^3+4m_\nu^2+m_\nu+4)C_5(\alpha_\nu)]\varepsilon_\nu^4\}+O(\varepsilon_\nu^2) \quad . \qquad (54)
\end{aligned}
$$

From Equation (27) we have

$$
\begin{aligned}
\frac{1}{\tilde{u}(z_\nu)} &= \frac{\tilde{p}'(z_\nu)}{\tilde{p}(z_\nu)} \quad , \\
&= \frac{m_\nu}{\varepsilon_\nu}+\tilde{S}_1(z_\nu) \quad . \qquad (55)
\end{aligned}
$$

The terms $\tilde{S}_k(z_\nu)$ for $k=2,3,\ldots$ are obtained from $\tilde{S}_1(z_\nu)$ using (9). It follows that

$$
\begin{aligned}
\tilde{S}_1(z_\nu) = {} & m_\nu\{C_2(\alpha_\nu)-[m_\nu C_2^2(\alpha_\nu)-2C_3(\alpha_\nu)]\varepsilon_\nu \\
& +[m_\nu^2C_2^3(\alpha_\nu)-3m_\nu C_2(\alpha_\nu)C_3(\alpha_\nu)+3C_4(\alpha_\nu)]\varepsilon_\nu^2 \\
& -[m_\nu^3C_2^4(\alpha_\nu)-4m_\nu^2C_2^2(\alpha_\nu)C_3(\alpha_\nu)+4m_\nu C_2(\alpha_\nu)C_4(\alpha_\nu) \\
& +2m_\nu C_3^2(\alpha_\nu)-4C_5(\alpha_\nu)]\varepsilon_\nu^3\}+O(\varepsilon_\nu^4) \quad , \\
\tilde{S}_2(z_\nu) = {} & m_\nu\{m_\nu C_2^2(\alpha_\nu)-2C_3(\alpha_\nu)-2[m_\nu^2C_2^3(\alpha_\nu)-3m_\nu C_2(\alpha_\nu)C_3(\alpha_\nu)+3C_4(\alpha_\nu)]\varepsilon_\nu \\
& +3[m_\nu^3C_2^4(\alpha_\nu)-4m_\nu^2C_2^2(\alpha_\nu)C_3(\alpha_\nu)+4m_\nu C_2(\alpha_\nu)C_4(\alpha_\nu)+2m_\nu C_3^2(\alpha_\nu) \\
& -4C_5(\alpha_\nu)]\varepsilon_\nu^2\}+O(\varepsilon_\nu^3) \quad , \\
\tilde{S}_3(z_\nu) = {} & m_\nu\{m_\nu^2C_2^3(\alpha_\nu)-3m_\nu C_2(\alpha_\nu)C_3(\alpha_\nu)+3C_4(\alpha_\nu) \\
& -3[m_\nu^3C_2^4(\alpha_\nu)-4m_\nu^2C_2^2(\alpha_\nu)C_3(\alpha_\nu)+4m_\nu C_2(\alpha_\nu)C_4(\alpha_\nu)+2m_\nu C_3^2(\alpha_\nu) \\
& -4C_5(\alpha_\nu)]\varepsilon_\nu\}+O(\varepsilon_\nu^2) \quad , \\
\tilde{S}_4(z_\nu) = {} & m_\nu\{m_\nu^3C_2^4(\alpha_\nu)-4m_\nu^2C_2^2(\alpha_\nu)C_3(\alpha_\nu)+4m_\nu C_2(\alpha_\nu)C_4(\alpha_\nu)+2m_\nu C_3^2(\alpha_\nu) \\
& -4C_5(\alpha_\nu)\}+O(\varepsilon_\nu) \quad . \qquad (56)
\end{aligned}
$$

Finally, we need an expansion of $T_k(z_\nu)$ about α_ν, namely $\tilde{T}_k(z_\nu)$. We first define the term $G_k(\alpha_\nu)$ by

$$
G_k(\alpha_\nu)=\sum_{i\neq\nu}\frac{m_i\varepsilon_i}{(\alpha_\nu-\alpha_i)^k}, \quad k=1,2,\ldots \quad .
$$

It is noted that $G_k(\alpha_\nu) = O(\varepsilon)$. It follows that

$$\begin{aligned}
\tilde{T}_k(z_\nu) &= \sum_{i\neq\nu} \frac{m_i}{(z_\nu - z_i)^k} \ , \\
&= \sum_{i\neq\nu} \frac{m_i}{(\alpha_\nu - \alpha_i + \varepsilon_\nu - \varepsilon_i)^k} \ , \\
&= \sum_{i\neq\nu} \frac{m_i}{(\alpha_\nu - \alpha_i)^k} \left[1 + \frac{\varepsilon_\nu - \varepsilon_i}{\alpha_\nu - \alpha_i}\right]^{-k} \ , \\
&= \sum_{i\neq\nu} \frac{m_i}{(\alpha_\nu - \alpha_i)^k} \left[1 - k\left(\frac{\varepsilon_\nu - \varepsilon_i}{\alpha_\nu - \alpha_i}\right) + O(\varepsilon^2)\right] \ , \\
&= \sum_{i\neq\nu} \frac{m_i}{(\alpha_\nu - \alpha_i)^k} - k\varepsilon_\nu \sum_{i\neq\nu} \frac{m_i}{(\alpha_\nu - \alpha_i)^{k+1}} \\
&\quad + k\sum_{i\neq\nu} \frac{m_i\varepsilon_i}{(\alpha_\nu - \alpha_i)^{k+1}} + O(\varepsilon^2) \ , \\
&= S_k(\alpha_\nu) - kS_{k+1}(\alpha_\nu)\varepsilon_\nu + kG_{k+1}(\alpha_\nu) + O(\varepsilon^2) \ . \qquad (57)
\end{aligned}$$

Note that $\nu = 1,\ 2,\ \ldots,\ N$ for each of the IFs described below.

4.2 One-point Iteration Functions

This section derives the AECs of our one-point IFs together with verification of their order of convergence.

4.2.1 Ernst Schröder's second-order IF

$$\begin{aligned}
\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Lambda_2(z_\nu) \ , \\
&= \varepsilon_\nu - m_\nu \tilde{u}(z_\nu) \ , \\
&= C_2(\alpha_\nu)\varepsilon_\nu^2 + O(\varepsilon_\nu^3) \ . \qquad (58)
\end{aligned}$$

The IF $z_\nu - \Psi_2(z_\nu)$ has the same AEC.

4.2.2 Joseph Traub's third-order IF

$$\begin{aligned}
\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Lambda_3(z_\nu) \ , \\
&= \varepsilon_\nu + m_\nu \frac{m_\nu - 3}{2}\tilde{u}(z_\nu) - m_\nu^2 \tilde{A}_2(z_\nu)\tilde{u}^2(z_\nu) \ , \\
&= [2^{-1}(m_\nu + 3)C_2^2(\alpha_\nu) - C_3(\alpha_\nu)]\varepsilon_\nu^3 + O(\varepsilon_\nu^4) \ . \qquad (59)
\end{aligned}$$

The IF $z_\nu - \Psi_3(z_\nu)$ has the same AEC.

4.2.3 Joseph Traub's fourth-order IF

$$\begin{aligned}\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Lambda_4(z_\nu)\ ,\\ &= \varepsilon_\nu - 6^{-1}m_\nu\left(m_\nu^2-6m_\nu+11\right)\tilde{u}(z_\nu)+m_\nu^2(m_\nu-2)\tilde{A}_2(z_\nu)\tilde{u}^2(z_\nu)\\ &\quad - m_\nu^3[2\tilde{A}_2^2(z_\nu)-\tilde{A}_3(z_\nu)]\tilde{u}^3(z_\nu)\ ,\\ &= \left[3^{-1}(m_\nu^2+6m_\nu+8)C_2^3(\alpha_\nu)-(m_\nu+4)C_2(\alpha_\nu)C_3(\alpha_\nu)+C_4(\alpha_\nu)\right]\varepsilon_\nu^4\\ &\quad + O(\varepsilon_\nu^5)\ . \end{aligned} \tag{60}$$

The IF $z_\nu - \Psi_4(z_\nu)$ has

$$\begin{aligned}\hat{\varepsilon}_\nu &= [3^{-1}(m_\nu^2+6m_\nu+5)C_2^3(\alpha_\nu)-(m_\nu+4)C_2(\alpha_\nu)C_3(\alpha_\nu)+C_4(\alpha_\nu)]\varepsilon_\nu^4\\ &\quad + O(\varepsilon_\nu^5).\end{aligned}$$

4.2.4 Joseph Traub's fifth-order IF

$$\begin{aligned}\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Lambda_5(z_\nu)\ ,\\ &= \varepsilon_\nu + 24^{-1}m_\nu\left(m_\nu^3-10m_\nu^2+35m_\nu-50\right)\tilde{u}(z_\nu)\\ &\quad - 12^{-1}m_\nu^2\left(7m_\nu^2-30m_\nu+35\right)\tilde{A}_2(z_\nu)\tilde{u}^2(z_\nu)\\ &\quad + 2^{-1}m_\nu^3(3m_\nu-5)[2\tilde{A}_2^2(z_\nu)-\tilde{A}_3(z_\nu)]\tilde{u}^3(z_\nu)\\ &\quad - m_\nu^4[5\tilde{A}_2^3(z_\nu)-5\tilde{A}_2(z_\nu)\tilde{A}_3(z_\nu)+\tilde{A}_4(z_\nu)]\tilde{u}^4(z_\nu)\ ,\\ &= [24^{-1}(6m_\nu^3+55m_\nu^2+150m_\nu+125)C_2^4(\alpha_\nu)\\ &\quad - 2^{-1}(2m_\nu^2+15m_\nu+25)C_2^2(\alpha_\nu)C_3(\alpha_\nu)+(m_\nu+5)C_2(\alpha_\nu)C_4(\alpha_\nu)\\ &\quad + 2^{-1}(m_\nu+5)C_3^2(\alpha_\nu)-C_5(\alpha_\nu)]\varepsilon_\nu^5+O(\varepsilon_\nu^6)\ . \end{aligned} \tag{61}$$

The IF $z_\nu - \Psi_5(z_\nu)$ has

$$\begin{aligned}\hat{\varepsilon}_\nu &= [24^{-1}(6m_\nu^3+55m_\nu^2+90m_\nu+41)C_2^4(\alpha_\nu)\\ &\quad - 2^{-1}(2m_\nu^2+15m_\nu+15)C_2^2(\alpha_\nu)C_3(\alpha_\nu)+(m_\nu+5)C_2(\alpha_\nu)C_4(\alpha_\nu)\\ &\quad + 2^{-1}(m_\nu+5)C_3^2(\alpha_\nu)-C_5(\alpha_\nu)]\varepsilon_\nu^5+O(\varepsilon_\nu^6).\end{aligned}$$

To the authors' best knowledge, this is the first time that these AECs have been explicitly given.

4.3 Simultaneous Iteration Functions

This section derives the AECs of our simultaneous IFs together with verifications of their orders of convergence. To obtain the simple zero versions, take the equations used in §3, and set

$$m_i = 1,\ \ i = 1,\ 2,\ \ldots, N,\ N = n\ . \tag{62}$$

4.3.1 Our third-order IF

$$
\begin{aligned}
\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Xi_3(z_\nu) \ , \\
&= \varepsilon_\nu - m_\nu \tilde{u}(z_\nu) - m_\nu \tilde{T}_1(z_\nu)\tilde{u}^2(z_\nu) \ , \\
&= C_2^2(\alpha_\nu)\varepsilon_\nu^3 - m_\nu^{-1} G_2(\alpha_\nu)\varepsilon_\nu^2 + O(\varepsilon^4) \ .
\end{aligned} \tag{63}
$$

The term $G_2(\alpha_\nu)$ is of order $O(\varepsilon)$, thus $\hat{\varepsilon}_\nu = O(\varepsilon^3)$.

4.3.2 Our fourth-order IF

$$
\begin{aligned}
\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Xi_4(z_\nu) \ , \\
&= [2^{-1}(3m_\nu + 5)C_2^3(\alpha_\nu) - 3C_2(\alpha_\nu)C_3(\alpha_\nu)]\varepsilon_\nu^4 - m_\nu^{-1} G_3(\alpha_\nu)\varepsilon_\nu^3 + O(\varepsilon^5) \ .
\end{aligned} \tag{64}
$$

The term $G_3(\alpha_\nu)$ is of order $O(\varepsilon)$, thus $\hat{\varepsilon}_\nu = O(\varepsilon^4)$.

4.3.3 Our fifth-order IF

$$
\begin{aligned}
\hat{\varepsilon}_\nu &= \varepsilon_\nu - \Xi_5(z_\nu) \ , \\
&= [6^{-1}(11m_\nu^2 + 36m_\nu + 31)C_2^4(\alpha_\nu) - 6(m_\nu + 2)C_2^2(\alpha_\nu)C_3(\alpha_\nu) \\
&\quad + 4C_2(\alpha_\nu)C_4(\alpha_\nu) + 2C_3^2(\alpha_\nu)]\varepsilon_\nu^5 - m_\nu^{-1} G_4(\alpha_\nu)\varepsilon_\nu^4 + O(\varepsilon^6) \ .
\end{aligned} \tag{65}
$$

The term $G_4(\alpha_\nu)$ is of order $O(\varepsilon)$, thus $\hat{\varepsilon}_\nu = O(\varepsilon^5)$.

5 R-order Convergence

Our simultaneous IFs, described in §3.3, utilise the function $T_k(z_\nu)$, given in Equation (33), to improve their efficiency over the one-point IFs, described in §3.2. In other words, only the old set of approximations, $\{z_i\}$, is used in computing the new set of approximations, $\{\hat{z}_i\}$, i.e. the new zeros are computed in a *parallel* fashion.

However, the new approximations can be improved by being computed in a *serial* fashion by rewriting $T_k(z_\nu)$ as

$$
\sum_{i<\nu} \frac{m_i}{(z_\nu - \hat{z}_i)^k} + \sum_{i>\nu} \frac{m_i}{(z_\nu - z_i)^k} \quad i = 1,\ 2,\ \ldots,\ N \ , \tag{66}
$$

where the new approximations are brought into play as soon as they become available. The formal term for the study of this technique is **R-order convergence**. See [AH74], [MP86], and [PR06] for further details.

However, one look at the AECs for our *simultaneous* IFs, see §4.3, will reveal that each of them contains two terms, and it is only the second of these that can be improved by using R-order convergence. This is therefore not an *advantage* in using our polynomial format, i.e. R-order convergence is not relevant in this case.

6 Overview of Stage 1

As noted in §1.1.1, the task in Stage 1 is firstly to find regions in the complex plane such that each region contains exactly one root of the polynomial $p(z)$ and secondly to find the multiplicity of each root of $p(z)$. The process to carry out this task is based on two theorems. One theorem, due to Marden [Mar66], provides a way of counting the number of roots of $p(z)$ in a disk. The other theorem, due to Lagouanelle [Lag66], provides a way of estimating the multiplicity of a root that is contained in a given small disk.

To start the process, a disk that contains all the roots of $p(z)$ is obtained. A suitable disk is defined by Dekker [Dek68]. This disk is centred at the origin and has the radius

$$2 \max_{0 \leq i \leq n-1} |a_i/a_n|^{1/(n-i)} \quad .$$

This disk is covered by a number of smaller disks, all with the same area. The roots in each smaller disk are counted, including multiplicities, using Marden's theorem. In detail, let $p^*(z)$ be the polynomial defined by

$$p^*(z) = \overline{a}_0 z^n + \overline{a}_1 z^{n-1} + \ldots + \overline{a}_n \quad ,$$

where $\overline{a}_i$ is the complex conjugate of a_i. Define the Schur transform of $p(z)$, $p(z) \mapsto T(p(z))$ by

$$T(p(z)) = \overline{a}_0 p(z) - a_n p^*(z) \quad .$$

The degree of $T(p(z))$ is strictly less than the degree of $p(z)$. Define the terms γ_i by

$$\gamma_i = T^i(p(z))\big|_{z=0}, i = 1, 2, \ldots, n \quad .$$

It can be shown that the γ_i are real numbers. Define the terms P_i by

$$P_i = \gamma_1 \gamma_2 \ldots \gamma_i, i = 1, 2 \ldots n \quad .$$

Marden's theorem states that if for some $k < n$, $P_k \neq 0$ but $T^{k+1}(p(z)) \equiv 0$, then $p(z)$ has $n-k$ zeros on the unit circle at the zeros of $T^k(p(z))$. If r of the P_i for $i = 1, 2, \ldots, k$ are negative, then $p(z)$ has r zeros inside the unit disk and $k-r$ zeros outside the unit disk. The theorem is the basis of a method for counting the zeros of $p(z)$ in any disk.

The covering disks that contain no roots are discarded. This process is iterated to produce a set of small disks, each of which contains at least one root of $p(z)$. When a disk is sufficiently small the multiplicity of a root in the disk can be estimated using Lagouanelle's theorem which states that the multiplicity m_i of a root α_ν of $p(z)$ is given by

$$m_i = \lim_{z \to \alpha_\nu} |u'(z)^{-1}|, i = 1, 2, \ldots, n \quad .$$

Any disk for which this multiplicity is equal to the number of roots in the disk is not reduced further. The final result is a set of disks such that each disk contains exactly one root of $p(z)$ with a known multiplicity.

The use of disks to constrain the locations of the zeros of $p(z)$ is inefficient because of the large overlaps among the smaller disks that are chosen to cover a larger disk. More efficient algorithms can be constructed using squares in place of disks. Further information can be found in [GM67] and [Far14].

7 Software and Experimental Results

The search stage (Stage 1) and all the IFs used in the second, iterative, stage are programmed in C using the GNU Multiple Precision Arithmetic Library (GMP) [Gra11]. The advantage of this library is that the precision of the arithmetic can be increased in order to remove errors due to a loss of precision. See [Far14] for the details, including computer programs and scripts in various languages. In addition, there is one Matlab program, `multiple.m`, which generates symbolic equations for our multiple IFs, both one-point and simultaneous, together with their orders of convergence and asymptotic error constants. A listing can be found in [Far14, pp. 132–139].

All of the formulae have been checked using Mathematica, version 10. The Matlab code and the Mathematica code can be provided by the authors on request.

The software was tested intensively on a database of 215 polynomials taken from the literature since 1950. Stage 1 was successful for all of the polynomials. Stage 2 succeeded on all but two polynomials in the database. The results showed that in general the higher order IFs converge faster than the lower order IFs. In more detail, the IFs (23), (24), (37) and the IFs

$$\hat{z}_\nu = z_\nu - m_\nu u(z_\nu)/(1 - T_1(z_\nu)u(z_\nu)), \tag{67}$$

$$\hat{z}_\nu = z_\nu - m_\nu u(z_\nu)/(2^{-1}(m_\nu + 1) - m_\nu A_2(z_\nu)u(z_\nu)), \tag{68}$$

$$\hat{z}_\nu = z_\nu - m_\nu u(z_\nu) - m_\nu \left(\sum_{i \neq \nu} m_i (z_\nu - (z_i - m_i u(z_i)))^{-1} \right) u^2(z_\nu) \tag{69}$$

are compared in [Far14]. The IF (67) is taken from [Ehr67], (68) is taken from [HP77] and (69) is the IFs Farmerv, as defined in [Far14]. These six IFs are implemented in parallel, and in addition, (37) and (67) are implemented in serial fashion, as described in §5. The performances of the eight IFs are measured by the number of iterations required for convergence. The results are shown in Table 7.2 in [Far14]. The IFs Farmerv outperforms the seven competing IFs as the degree of the polynomial is increased.

The database contains the following two extreme polynomials. The Wilkinson polynomial [Wil59]

$$p(z) = \prod_{i=1}^{20} (z + i) \tag{70}$$

has extremely large coefficients, but each root has multiplicity one. In contrast, the polynomial [BF00]

$$p(z) = (z^4 - 16^{-1})^{40}(z^4 - (2^{-1} + \eta)^4), \eta = 4096^{-1} \quad , \tag{71}$$

has four roots of multiplicity 40 and each of these roots is near to a root of multiplicity one. The roots of (70) were found to a high accuracy in Stage 2 using just two iterations of Farmerv. The roots of (71) were found to a high accuracy by each of the eight IFs using one iteration.

8 Conclusion

This paper demonstrates a straightforward mechanism for deriving one-point IFs of order two or more, or simultaneous IFs of order three or more. In addition, by taking the most general approach, i.e. multiple zeros, IFs for polynomials with only simple zeros are just a special case.

These IFs have been extensively tested computationally. In the context of our global algorithm, outlined in §1.1, we have tested polynomials of high degree (up to degree 400) and polynomials with zeros of high multiplicity (up to order 40), with complete success, i.e. they converge to within the accuracy that is required.

9 Bibliography

[AH74] Götz Alefeld and Jürgen Herzberger. On the convergence speed of some algorithms for the simultaneous approximation of polynomial zeros. *SIAM Journal of Numerical Analysis*, 11:237–243, 1974.

[BF00] Dario Bini and Giuseppe Fiorentino. Numerical computation of polynomial roots using MPSolve version 2.2. Technical report, Dipartimento di Matematica, Università di Pisa, Via Buonarroti 2, 56127 Pisa, January 2000.

[Dek68] Theodorus Dekker. Newton-Laguerre iteration. In *Colloque International du CNRS, Programmation en Mathématiques Numériques*, volume 165, pages 189–200, Besançon, France, 1968.

[Ehr67] Louis Ehrlich. A modified Newton method for polynomials. *Communications of the ACM*, 10(2):107–108, 1967.

[Far13] Michael Farmer. *Computing the Zeros of Polynomials using the Divide and Conquer Approach*. PhD thesis, Department of Computer Science and Information Systems, Birkbeck, University of London, England, September 2013.

[Far14] Michael Farmer. Computing the zeros of polynomials using the divide and conquer approach. Website, 2014. `http://www.dcs.bbk.ac.uk/research/recentphds/farmer.pdf`.

[FL75] Michael Farmer and George Loizou. A class of iteration functions for improving, simultaneously, approximations to the zeros of a polynomial. *BIT*, 15:250–258, 1975.

[FL77] Michael Farmer and George Loizou. An algorithm for the total, or partial, factorization of a polynomial. *Mathematical Proceedings of the Cambridge Philosophical Society*, 82(427):427–437, 1977.

[FL85] Michael Farmer and George Loizou. Locating multiple zeros interactively. *Computers and Mathematics with Applications*, 11(6):595–603, 1985.

[GM67] Irene Gargantini and W Münzner. An experimental program for the simultaneous determination of all zeros of a polynomial. Technical Report RZ-237, IBM Zurich Research Laboratory, 8803 Rüschlikon-ZH, Switzerland, April 1967.

[Gra11] Törbjorn Granlund. *GNU Multiple Precision Arithmetic Library (GMP)*. Free Software Foundation, 51 Franklin Street, Boston, MA 02110-1301 USA, 5.0.2 edition, May 2011.

[Hen74] Peter Henrici. *Applied and Computational Complex Analysis*, volume 1 of *Pure and Applied Mathematics*. Wiley-Interscience, 1974. Contents: Power Series—Integration—Conformal Mapping—Location of Zeros.

[Hou70] Alston Householder. *The Numerical Treatment of a Single Nonlinear Equation*. McGraw-Hill Inc., New York, June 1970.

[HP77] Eldon Hansen and Merrell Patrick. A family of root finding methods. *Numerische Mathematik*, 27:257–269, 1977.

[Kis54] Ignác Kiss. Über eine Verallgemeinerung des Newtonschen Nǟherungsverfahrens. *Zeitschrift für Agnewandte Mathematik und Mechanik*, 34(1):68–69, January/February 1954.

[Lag66] Jean-Louis Lagouanelle. Sur une méthode de calcul de l'ordre de multiplicité des zéros d'un polynôme. *Comptes Rendus de l'Académie des Sciences, Série A*, 262(11):626–627, March 1966.

[Mar66] Morris Marden. *Geometry of Polynomials (Mathematical Surveys Number 3)*. American Mathematical Society, Providence, Rhode Island, 2nd edition, 1966.

[MP86] Gradimir Milovanović and Miodrag Petković. On computational efficiency of the iterative methods for the simultaneous approximation of polynomial zeros. *ACM Transactions on Mathematical Software*, 12(4):295–306, December 1986.

[Pet89] Miodrag Petković. Iterative methods for simultaneous inclusion of polynomial zeros. In A Dold and B Eckmann, editors, *Lecture Notes in Mathematics*, volume 1387. Springer-Verlag, Berlin, 1989.

[PPŽ03] Ljiljana Petković, Miodrag Petković, and Dragan Živković. Hansen-Patrick's family is of Laguerre type 1. *Novi Sad Journal of Mathematics*, 33(1):109–115, 2003.

[PR06] Miodrag Petković and Lidija Rančić. A family of root-finding methods with accelerated convergence. *Computers and Mathematics with Applications*, 51:999–1010, 2006.

[Sch70] Ernst Schröder. Über unendlich viele Algorithmen zur Auflösung der Gleichungen. *Mathematische Annalen*, 2:317–365, 1870.

[Tra82] Joseph Traub. *Iterative Methods for the Solution of Equations*. Chelsea Publishing Company, New York, USA, 1982.

[Voi71] Robert Voigt. Orders of convergence for iterative procedures. *SIAM Journal on Numerical Analysis*, 8(2):222–243, 1971.

[Wil59] James Wilkinson. The evaluation of the zeros of ill-conditioned polynomials. Part I. *Numerische Mathematik*, 1:150–166, 1959.