2009 American Control Conference
Hyatt Regency Riverfront, St. Louis, MO, USA
June 10-12, 2009

Derivative-Free Family of Higher Order Root Finding Methods

Mohammed A. Hasan
Department of Electrical & Computer Engineering
University of Minnesota Duluth
E.mail:mhasan@d.umn.edu

Abstract: *Most higher order root finding methods require evaluation of a function and/or its derivatives at one or multiple points. There are cases where the derivatives of a given function are costly to compute. In this paper, higher order methods which do not require computation of any derivatives are derived. Asymptotic analysis has shown that these methods are approximations of root iterations. One of the main features of the proposed approaches is that one can develop multi-point derivative-free methods of any desired order. For lower order methods, these correspond to the Newton, and Ostrowski iterations. Several examples involving polynomials and entire functions have shown that the proposed methods can be applied to polynomial and non-polynomial equations.*

Keywords: Zeros of polynomials, Zeros of analytic functions, derivative free methods, Root iterations, root-finding, order of convergence, Halley's Method, Newton's Method, Square root iteration, higher order methods, Ostrowski method

1 Introduction

Let f be a polynomial of degree n with coefficients in $\mathcal{C}$, where $\mathcal{C}$ denotes the field of complex numbers, and assume that the zeros of f are $\xi_1, \cdots, \xi_n$. There are many higher order methods for computing zeros of the polynomial f. These methods may be classified into one-point and multi-point zero-finding methods. In one-point zero-finding methods, new approximations in each iteration are found by using the values of f and perhaps its derivatives at only one point. In multi-point methods, new approximations are obtained by using the values of f and sometimes its derivatives at a number of points. Newton's and Halley's methods are examples of one-point methods, while the secant and Muller's methods are examples of multi-point methods. Analysis related to one-point zero-finding methods appears in [1]-[2], while multi-point methods are analyzed in [3]-[4]. Good treatments of general root-finding methods can be found in [5]-[6] and the references therein.

In this paper, we will analyze some known methods and convert them into derivative-free methods. This conversion is based on optimal approximation of derivatives using multi-point computation of the original function. Other methods proposed in this work are based on approximating the first and higher order derivatives of $\log(f(z))$, the natural logarithm of $f(z)$. These approximations are then utilized for developing derivative-free multi-point root iteration methods.

The following notation will be used throughout. The sets $\mathbb{R}$ and $\mathcal{C}$ denote the fields of real and complex numbers, respectively. If $z \in \mathcal{C}$, then $z = x + jy$ where $x, y \in \mathbb{R}$ and $j = \sqrt{-1}$. The number $z^* = x - jy$ is the complex conjugate of z. In this presentation, it will be assumed that f is a polynomial of degree n with simple zeros $\xi_1, \cdots, \xi_n$, unless stated otherwise.

For a given algorithm, the order of convergence is defined as follows: Let z_k be a sequence of complex numbers and $\lambda \in \mathcal{C}$. If there is a real number $r \geq 1$ and a constant $C_r \in \mathbb{R}$, such that $|z_{k+1} - \lambda| \leq C_r |z_k - \lambda|^r$ as $k \to \infty$ whenever z_0 is sufficiently near λ, then the sequence z_k is said to be order r convergent to λ. If $r = 1$, we further require that $C_r < 1$ and we call C_r the asymptotic linear convergence constant for the sequence if it is the smallest such a constant. Alternatively, assume that the sequence z_k is generated by the fixed point iteration $z_{k+1} = \Phi(z_k)$ where Φ is analytic in a bounded neighborhood V_r of a root ξ of a polynomial f having only simple roots. If for some ξ we have $\Phi(\xi) = \xi$, $\Phi'(\xi) = 0, \cdots, \Phi^{(r-1)}(\xi) = 0$ and $\Phi^{(r)}(\xi) \neq 0$, then the root-finding algorithm is at least rth order convergent. Here, $\Phi'(z), \Phi''(z), \Phi'''(z), \cdots, \Phi^{(r)}(z)$, denote the first, second, third, and rth derivatives of Φ evaluated at the complex number z. We also use the convention that $\Phi^{(k)} = \Phi$ if $k = 0$.

Most multi-point iterations have fractional order of convergence, however it is often the case that one-point methods have integer order of convergence. The following result provides conditions for a given one-point iteration to be of a given order.

Theorem 1[7]. *Let f be a polynomial of degree n with zeros $\xi_1, \cdots, \xi_n$. Let g be analytic function in a neighborhood of ξ_k, $k = 1, \cdots, n$. Assume that each zero of f is simple. Then the iteration $\Phi(z) = z - \frac{f(z)}{g(z)}$ is rth order iff $g^{(i)}(\xi_j) = \frac{f^{(i+1)}(\xi_j)}{i+1}$ for $j = 1, \cdots, n$ and $i = 0, \cdots, r-1$. Hence if ξ is a simple zero of f, then the Taylor expansion of g around $z = \xi$ is given as:*

$$g(z) = \sum_{k=0}^{r-1} \frac{f^{(k+1)}(\xi)}{(k+1)!}(z-\xi)^k + O((z-\xi)^r) \qquad (1a)$$

Additionally, a method is of infinite order if

$$g(z) = \sum_{k=0}^{\infty} \frac{f^{(k+1)}(\xi)}{(k+1)!}(z-\xi)^k = \frac{f(z)}{z-\xi}. \qquad (1b)$$

Hence if g can be expressed as

$$g = f' + \sum_{k=1}^{r-1} h_k f^k + O(f^r), \qquad (1c)$$

where $\{h_k\}_{k=1}^{r-1}$ are analytic functions around neighborhoods of the zeros of f, then Φ is at least rth order fixed point function.

Proof. A version of this result is stated in [7]. The proof follows by showing that $\Phi(\xi) = \xi$, and $\Phi^{(k)}(\xi) = 0$, for $k = 1, \cdots, r-1$.

2 Review of Some Zero-Finding Methods

To understand the derivation and convergence behavior of different methods, a brief review of well-known methods [8]-[12] is given in this section. Newton's method and many of its variations are quadratically convergent, while Halley's, Chebyshev, the square root iteration, Laguerre, and Euler methods are cubically convergent. Derivations and new perspective regarding some of these methods, and other multi-point methods are highlighted.

The Secant Method: This is a two-point iteration and is given by

$$z_{k+1} = z_k - \frac{f(z_k)(z_k - z_{k-1})}{f(z_k) - f(z_{k-1})}. \tag{2}$$

This corresponds to interpolating the equation $y = f(z)$ using the straight line between the points $(z_k, f(z_k))$, $(z_{k-1}, f(z_{k-1}))$, where z_{k-1} and z_k are approximations of a zero of f. The next iterate is now given by the root of the equation $y = 0$. This yields the recurrence relation (2).

Muller's Method: In this method, three points are used to locally fit a parabola to approximate a function f. New approximations are obtained from the intersection of the parabola with the z-axis. The three initial values needed are denoted as z_k, z_{k-1} and z_{k-2}. The parabola passes through the three points $(z_k, f(z_k))$, $(z_{k-1}, f(z_{k-1}))$ and $(z_{k-2}, f(z_{k-2}))$, may be written as $y = az^2 + bz + c$, for some $a, b, c \in \mathcal{C}$. The next iterate is now given by the root of a quadratic equation $y = 0$.

Newton's Method: Newton's method and many of its variations are quadratically convergent. It is derived from the Taylor expansion of f: $f(z_k) = f(\xi) + (z_k - \xi)f'(z_k) + O(z_k - \xi)^2$, where z_k is an approximation of a zero ξ of f. The iteration formula for the Newton's method is

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k)}. \tag{3}$$

Note that the secant method follows from Newton's method by replacing $f'(z_k)$ with the quotient $\frac{f(z_k)-f(z_{k-1})}{z_k - z_{k-1}}$.

Halley's Method: This is a cubically convergent method for computing simple zeros of f. The iteration formula for the Halley's method is

$$z_{k+1} = z_k - \frac{f(z_k)}{f'(z_k) - \frac{f(z_k)f''(z_k)}{2f'(z_k)}}. \tag{4}$$

There are many approaches in the literature for the derivation of Halley's method. As shown in [13], Halley's method can be obtained by applying Newton's method to the function $\frac{f}{\sqrt{f'}}$. Gerlach [14], gives a generalization of this approach.

It is interesting to note that Halley's method can also be derived using the generalized Taylor expansion. The first few terms of the generalized Taylor expansion [15] is given by

$$f(z) = f(a) + \frac{z-a}{2}(f'(a) + f'(z)) - \frac{(z-a)^3}{12}f'''(\theta). \tag{5}$$

where θ is between a and z. Hence, if $a = \xi$ is a zero of f, i.e.,$f(\xi) = 0$, then

$$\xi \approx \Phi(z) = z - \frac{2f(z)}{f'(\xi) + f'(z)}. \tag{6a}$$

Theorem 1 may be used to show that this implicit iteration is third order. Clearly, if it is assumed that $f(\xi) = 0$ and $f'(\xi) = 1$, the following third order iteration will be obtained

$$\Phi(z) = z - \frac{2f(z)}{1 + f'(z)}. \tag{6b}$$

Generally, $f'(\xi) \neq 1$ and thus we may consider the function $g = \frac{f}{f'}$. Then $g' = \frac{f'^2 - ff''}{f'^2}$, i.e., $g'(\xi) = 1$ provided that ξ is a simple zero of f. Consequently, the iteration function (6b) for solving $g(z) = 0$ simplifies to

$$\Phi(z) = z - \frac{2\frac{f(z)}{f'(z)}}{1 + \frac{f'(z)^2 - f(z)f''(z)}{f'(z)^2}} = z - \frac{f(z)}{f'(z) - \frac{f(z)f''(z)}{2f'(z)}} \tag{6c}$$

which is the Halley's iteration. Up to the author knowledge these derivations of Halley's method using the generalized Taylor expansion are not discussed in the literature.

Another version of Halley's iteration also follows from (6a) by replacing $f'(\xi)$ with $h(z) = f'(z - \frac{f(z)}{f'(z)})$, where $z - \frac{f(z)}{f'(z)}$ is the Newton's approximation of ξ. Clearly the first few terms of the Taylor expansion of h around z is $h(z) = f'(z) - \frac{f(z)f''(z)}{f'(z)} + O(f(z)^2)$.

Ostrowski's and Root Iteratons: Ostrowski's method is another cubically convergent iteration and is also known as the square root iteration [5]. The square root iteration has been generalized in [16] to obtain radical methods of any desired order.

Root iterations are based on the observation that if $f(z) = \Pi_{k=1}^{n}(z - \xi_k)$, then

$$\{\frac{f'(z)}{f(z)}\}^{(r)} = r!(-1)^r \sum_{k=1}^{n} \frac{1}{(z-\xi_k)^{r+1}}. \tag{7}$$

Thus if $r = 1$ and z is sufficiently close to a zero ξ of f, then higher order logarithmic derivatives of f can be expressed as:

$$\{\frac{f'(z)}{f(z)}\}' = -\sum_{k=1}^{n} \frac{1}{(z-\xi_k)^2} \approx \frac{-1}{(z-\xi)^2}, \tag{8}$$

or equivalently

$$\frac{f'(z)^2 - f(z)f''(z)}{f(z)^2} \approx \frac{1}{(z-\xi)^2}. \tag{9}$$

By solving (9) for ξ and setting $z_{k+1} = \xi$, it follows that

$$z_{k+1} = z_k - \frac{f(z_k)}{\sqrt{f'^2(z_k) - f(z_k)f''(z_k)}}, \tag{10}$$

which is the square root iteration.

Generally, if z is close to ξ, (7) implies that

$$\frac{1}{r!}(-1)^r\{\frac{f'(z)}{f(z)}\}^{(r)} \approx \frac{1}{(z-\xi)^{r+1}}. \tag{11}$$

From (11), an $(r+1)$ root iteration can be expressed as

$$\Phi(z) = z - \frac{1}{\sqrt[r+1]{\frac{1}{r!}(-1)^r\{\frac{f'(z)}{f(z)}\}^{(r)}}}. \tag{12}$$

For $r = 2$, (12) simplifies to

$$\phi(z) = z - \frac{f(z)}{\sqrt[3]{f'(z)^3 - \frac{3}{2}f(z)f'(z)f''(z) + \frac{1}{2}f(z)^2 f'''(z)}}, \tag{13}$$

which is a fourth order iteration [16].

An Example: Consider the entire function $f(z) = \sin(z)$, then $f'(z) = \cos(z)$, $f''(z) = -\sin(z)$, $f(0) = 0$ and $f'(0) = 1$. By applying (6b) we obtain

$$\begin{aligned} z_{k+1} &= z_k - \frac{2\sin(z_k)}{1+\cos(z_k)}, \\ &= z_k - \frac{2\sin(\frac{z_k}{2})\cos(\frac{z_k}{2})}{\cos^2(\frac{z_k}{2})}, \\ &= z_k - \frac{2\sin(\frac{z_k}{2})}{\cos(\frac{z_k}{2})}, \\ &= z_k - 2\tan(\frac{z_k}{2}). \end{aligned}$$

Note that if the Halley's method is applied, the following iteration will follow

$$z_{k+1} = z_k - \cos(z_k)\sin(z_k) = z_k - \frac{1}{2}\sin(2z_k).$$

Using the square root iteration, we obtain

$$z_{k+1} = z_k - \sin(z_k).$$

This shows that there are many ways of developing third order iterations for the same equation.

3 Derivative-Free Higher Order Methods

Let $h \in \mathcal{C}$ such that $h \neq 0$ and $|h|$ is sufficiently small. Let f be a polynomial of degree n and consider the function F_2 defined as

$$F_2(z,h) = f(z+h)f(z-h) - f(z)^2. \tag{14}$$

Clearly, F_2 is an even function of h and hence it can easily be verified that $F_2(x,h) = G_2(z,h^2)$ for some function G_2. From the Taylor expansion of $f(z+h)$ and $f(z-h)$ around z it follows that

$$\begin{aligned} F_2(z,h) &= f(z+h)f(z-h) - f(z)^2 \\ &= -h^2(f'(z)^2 - f(z)f''(z)) + O(h^4), \end{aligned} \tag{15}$$

and therefore, one can show that

$$\frac{-h^2 f(z)^2}{F_2(z,h)} = \frac{f(z)^2}{f'(z)^2 - f(z)f''(z)} + O(h^2). \tag{16}$$

Now considering the expression $\frac{f^2}{f'^2 - ff''}$ and comparing that with the term in Ostrowski method, we obtain

$$\phi(z) = z - \frac{hf(z)}{\sqrt[2]{f(z)^2 - f(z+h)f(z-h)}}, \tag{17}$$

which is an approximated square root iteration. Specifically, it can be shown that (17) is asymptotically of order 3 as $h \to 0$. The main advantages of this iteration is that it only requires function computation at three points $z, z+h, z-h$, and without calculating any derivatives.

An asymptotically fourth order method that does not involve computation of derivatives may be derived as follows. Let $F_3(z,h) = f(z+h)f(z+wh)f(z+w^2h) - f(z)^3$, where w is a primitive cube root of 1, i.e., $\omega = \frac{-1+j\sqrt{3}}{2}$, or $\omega = \frac{-1-j\sqrt{3}}{2}$. It is easy to verify that $F_3(z,h) = F_3(z,wh) = F_3(z,w^2h)$. This implies that $F_3(z,h) = G_3(z,h^3)$ for some function G_3. Using this symmetric property and after algebraic simplifications, the expression $F_3(z,h) = f(z+h)f(z+wh)f(z+w^2h) - f(z)^3$ can be written as

$$F_3(z,h) = h^3(f'(z)^3 - \frac{3}{2}f(z)f'(z)f''(z) + \frac{1}{2}f(z)^2 f'''(z)) + O(h^6) \tag{18}$$

or $F_3(z,h) = \frac{1}{2}f(z)^3(\frac{f'(z)}{f(z)})''h^3 + O(h^6)$. Thus using the root iteration formula (12) with $r=3$, we obtain

$$\phi(z) = z - \frac{hf(z)}{\sqrt[3]{f(z+h)f(z+wh)f(z+w^2h) - f(z)^3}}, \tag{19}$$

which is asymptotically a fourth order iteration near a simple zero of f.

A generalization of the above observations to multi-point version of the root iteration is given in the next result.

Theorem 2. *Let w be a primitive r-th root of 1, and consider the following function*

$$F_r(z,h) = f(z+h)f(z+wh)f(z+w^2h)\cdots f(z+w^{r-1}h) - f(z)^r, \tag{20a}$$

Then $F_r(z,h)$ can be written as

$$\begin{aligned} &F_r(z,h) = f(z)^r - h^r\{\sum (z-z_{k_1})^r \cdots (z-z_{k_{n-1}})^r\} \\ &+ h^{2r}\{\sum (z-z_{k_1})^r \cdots (z-z_{k_{n-2}})^r\} + \cdots \\ &h^{r(n-1)}\{\sum (z-z_k)^r\} + (-1)^n h^{rn}, \end{aligned} \tag{20b}$$

where $\{k_1, \cdots, k_r\}$ is an rth order combination of the set of integers $\{1, 2, \cdots, n\}$.

Proof: Assume that $f(z) = (z-\xi_1)(z-\xi_1)\cdots(z-\xi_n)$, then for each $1 \le k \le n$ there holds:

$$f(z+w^k h) = (z+w^k h - \xi_1)(z+w^k h - \xi_1)\cdots(z+w^k h - \xi_n).$$

Therefore,

$$F_r(z,h) = \Pi_{k=1}^n\{(z-\xi_k)^r - h^r\} - f(z)^r,$$

from which (20b) follows. Clearly, $F_r(z,h) = F_r(z,wh) = F_r(z,w^2h)\cdots = F_r(z,w^{r-1}h)$. Consequently, $F_r(z,h) = G_r(z,h^r)$ for some function G_r. This implies that

$$\begin{aligned} \frac{F_r(z,h)}{f(z)^r} &= \frac{\Pi_{k=1}^n\{(z-\xi_k)^r - h^r\} - f(z)^r}{f(z)^r} \\ &= -h^r \sum_{k=1}^n \frac{1}{(z-\xi_k)^r} + O(h^{2r}), \end{aligned}$$

and hence

$$\pm \sum_{k=1}^n \frac{1}{(z-\xi_k)^r} = \frac{F_r(z,h) - f(z)^r}{h^r f(z)^r} + O(h^r).$$

The $\pm$ depends on whether r is even or odd.

Recall that if $f(z) = \Pi_{k=1}^n(z-\xi_k)$, then

$$\{\frac{f'}{f}\}^{(r)} = r!(-1)^r \sum_{k=1}^n \frac{1}{(z-\xi_k)^{r+1}}.$$

This shows that

$$\{\frac{f'}{f}\}^{(r)} = K_r \sum_{k=1}^n \frac{1}{(z-\xi_k)^r} = \frac{F_r(z,h) - f(z)^r}{h^r f(z)^r} + O(h^r), \tag{21}$$

for some constant K_r which is a function of r.

4 Derivative Approximations

The first order derivative of f may be approximated using forward, backward, or central differences. It is known that forward, backward differences are of order $O(\epsilon)$ while the central difference approximation is of order $O(\epsilon^2)$. Specifically, forward, backward, and central differences are respectively given by:

$$\frac{f(z)-f(z-\epsilon)}{\epsilon}=f'(z)+O(\epsilon), \tag{22a}$$

$$\frac{f(z+\epsilon)-f(z)}{\epsilon}=f'(z)+O(\epsilon), \tag{22b}$$

$$\frac{f(z+\epsilon)-f(z-\epsilon)}{2\epsilon}=f'(z)+O(\epsilon^2). \tag{23c}$$

4.1 Two-Point Approximation

For two-point approximation of $f'(z)$, it can be shown that the approximation

$$f'(z)\approx\frac{f(z+\epsilon)-f(z-\epsilon)}{2\epsilon}, \tag{23d}$$

is optimal in the sense that if

$$f'(z)=\alpha_1 f(z+h_1\epsilon)+\alpha_2 f(z+h_2\epsilon)+O(\epsilon^2)$$

for some nonzero numbers $\alpha_1,\alpha_2,h_1,h_2,h_1\neq h_2$, then $\alpha_1=-\alpha_2$ and $h_1=-h_2$. By using the Taylor expansions of $f(z+h_1\epsilon)$ and $f(z+h_2\epsilon)$ around z, it follows that

$$\begin{aligned}&\alpha_1+\alpha_2=0,\\&\alpha_1h_1+\alpha_2h_2=1,\\&\alpha_1h_1^2+\alpha_2h_2^2=0.\end{aligned} \tag{24}$$

The sysyem of equations in (24)implies that

$$\begin{aligned}&\alpha_2=-\alpha_1,\\&h_1^2=h_2^2,\\&\alpha_1(h_1-h_2)=1.\end{aligned} \tag{25}$$

Therefore, acceptible solutions are $h_2=-h_1,\alpha_2=-\alpha_1=-\frac{1}{2h_1}$. Note that we can not obtain a higher order two-point approximation of the form

$$f'(z)=\alpha_1 f(z_k+h_1\epsilon)+\alpha_2 f(z+h_2\epsilon)+O(\epsilon^r)$$

where r is a positive integer such that $r\geq 3$. In this case, the following equation must hold (if $r=3$):

$$\alpha_1h_1^3+\alpha_2h_2^3=0. \tag{26}$$

This equation along with the third equation of (25) imply that $\alpha_2h_2^2(h_1-h_2)=0$ which yields $\alpha_2=0$, $h_2=0$, or $h_1=h_2$. Each of these solutions is unacceptible since they contradict the solutions of the three equations in (24). This shows that optimal two-point approximation of f' is second order. From this observation one may assume that

$$\alpha_1h_1^3+\alpha_2h_2^3=\gamma, \tag{27}$$

where γ is any nonzero complex number. By incorporating (27) into (24), we obtain $2\alpha_1h_1^3=\gamma$ and $2\alpha_1h_1=1$. These two equations show that $h_1^2=h_2^2=\gamma$. Thus $h_1=-h_2=\frac{1}{\sqrt{\gamma}}$ and $\alpha_1=-\alpha_2=\frac{1}{2\sqrt{\gamma}}$. Consequently, a second order approximation of f' has the form

$$f'(z)=\frac{f(z+\sqrt{\gamma}\epsilon)-f(z-\sqrt{\gamma}\epsilon)}{2\sqrt{\gamma}\epsilon}+O(\epsilon^2). \tag{28}$$

An alternative method for determining h_1 and h_2 is by noting that h_1 and h_2 must satisfy a second order polynomial equation $h^2+a_1h+a_2=0$ where a_1,a_2 are obtained as:

$$\begin{bmatrix}0&1\\1&0\end{bmatrix}^{-1}\begin{bmatrix}0\\\gamma\end{bmatrix}=\begin{bmatrix}-a_2\\-a_1\end{bmatrix} \tag{29}$$

where $\gamma=\alpha_1h_1^3+\alpha_2h_2^3$. Thus $\begin{bmatrix}-a_2\\-a_1\end{bmatrix}=\begin{bmatrix}\gamma\\0\end{bmatrix}$. This shows that h_1,h_2 satisfy the second order equation $h^2-\gamma=0$. The last equation has two solutions $h_1=-h_2=\frac{1}{\sqrt{\gamma}}$ and α_1, α_2 can be determined from solving the following system

$$\begin{bmatrix}1&1\\h_1&h_2\end{bmatrix}\begin{bmatrix}\alpha_1\\\alpha_2\end{bmatrix}=\begin{bmatrix}0\\1\end{bmatrix}. \tag{30}$$

Note that exact derivatives will be obtained if (28) is applied to polynomials of degree at most 2.

4.2 Three-Point Approximation

The second approach can be easily generalized for three-point approximation of first order derivative as follows. Let a three-point approximation be given by

$$f'(z)=\sum_{k=1}^{3}\alpha_k f(z+h_k\epsilon)+O(\epsilon^r). \tag{31}$$

The objective is to find h_1,h_2,h_3 and $\alpha_1,\alpha_2,\alpha_3$, and the highest integer r such that the h_k's are distinct and each of α_k is nonzero. The Taylor expansion of (31) around z leads to the following equations:

$$\begin{aligned}&\alpha_1+\alpha_2+\alpha_3=0,\\&\alpha_1h_1+\alpha_2h_2+\alpha_3h_3=1,\\&\alpha_1h_1^2+\alpha_2h_2^2+\alpha_3h_3^2=0,\\&\alpha_1h_1^3+\alpha_2h_2^3+\alpha_3h_3^3=0,\\&\alpha_1h_1^4+\alpha_2h_2^4+\alpha_3h_3^4=\gamma_1,\\&\alpha_1h_1^5+\alpha_2h_2^5+\alpha_3h_3^5=\gamma_2.\end{aligned} \tag{32}$$

The optimal order of approximation can be determined first since the case where $\gamma_1=0$ implies that

$$H=\begin{bmatrix}h_1^2&h_2^2&h_3^2\\h_1^3&h_2^3&h_3^3\\h_1^4&h_2^4&h_3^4\end{bmatrix}$$

is singular. Since the determinant $|H|$ is given by

$$|H|=h_1^2h_2^2h_3^2(h_2-h_1)(h_3-h_2)(h_3-h_1),$$

it follows that H is nonsingular if and only if h_1,h_2,h_3 are nonzero distinct numbers. This shows that γ_1 can not be zero and consequently $r=3$. Now assuming that $r=3$ and γ_1,γ_2 are arbitrary complex numbers such that $\gamma_1\neq 0$, then h_1,h_2,h_3 must solve the equation $h^3+a_1h^2+a_2h+a_3=0$, where a_1,a_2,a_3 are determined from the following equation

$$\begin{bmatrix}0&1&0\\1&0&0\\0&0&\gamma\end{bmatrix}^{-1}\begin{bmatrix}0\\\gamma_1\\\gamma_2\end{bmatrix}=\begin{bmatrix}-a_3\\-a_2\\-a_1\end{bmatrix}, \tag{33}$$

or

$$\begin{bmatrix}a_3\\a_2\\a_1\end{bmatrix}=-\begin{bmatrix}0&1&0\\1&0&0\\0&0&\frac{1}{\gamma_1}\end{bmatrix}\begin{bmatrix}0\\\gamma_1\\\gamma_2\end{bmatrix}=-\begin{bmatrix}\gamma_1\\0\\\frac{\gamma_2}{\gamma_1}\end{bmatrix}. \tag{34}$$

This shows that h_1,h_2,h_3 satisfy the equation

$$h^3-\frac{\gamma_2}{\gamma_1}h^2-\gamma_1=0. \tag{35}$$

Since γ_1,γ_2 are arbitrary complex numbers such that $\gamma_1\neq 0$, one may assume for convenience that $\gamma_2=0$, and $\gamma_1=\gamma^3$ for some $\gamma\neq 0$, then h_1,h_2,h_3 satisfy the third order equation $h^3-\gamma^3=0$, i.e, $h=\gamma,w\gamma,w^2\gamma$, where w is a primitive cube root of 1, (i.e., $w^3=1$, and $w\neq 1$). Thus the corresponding values of α_1,α_2 and α_3 are determined from the system

$$\begin{bmatrix} 1 & 1 & 1 \\ h_1 & h_2 & h_3 \\ h_1^2 & h_2^2 & h_3^2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}. \tag{36}$$

Hence

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ h_1 & h_2 & h_3 \\ h_1^2 & h_2^2 & h_3^2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{3\gamma} \\ \frac{w^2}{3\gamma} \\ \frac{w}{3\gamma} \end{bmatrix}. \tag{37}$$

Consequently a third order approximation for $f'(z)$ is given by

$$\begin{aligned} f'(z) &= \frac{1}{3\gamma} f(z+\gamma\epsilon) + \frac{w^2}{3\gamma} f(z+\gamma w\epsilon) \\ &+ \frac{w}{3\gamma} f(z+\gamma w^2\epsilon) + O(\epsilon^3). \end{aligned} \tag{38}$$

As indicated earlier, this approximation is not unique in that it depends on the parameter γ. Generally one may obtain a two parameter approximation of f' by assigning different values of γ_1 and γ_2.

Remark: The cubic equation (35) can be solved using Cardano's formula [17] or any other cubic equation solver. However, the main characterstics of (35) is that if h_1, h_2, h_3 are its zeros then $h_1h_2 + h_1h_3 + h_2h_3 = 0$. Thus one may choose γ_1 and γ_2 so as to guarantee that h_1, h_2, h_3 are real. Choosing real values of h_1, h_2, h_3 makes it more convenient and more efficient for approximating derivatives using real arithmetic. Thus, assuming that h_1 and h_2 are real numbers such that $h_1+h_2 \neq 0$, then $h_3 = \frac{-h_1h_2}{h_1+h_2}$. This implies that $\gamma_1 = \frac{-h_1^2h_2^2}{h_1+h_2}$ and $\gamma_2 = \frac{h_1^2+h_1h_2+h_2^2}{h_1+h_2}$.

4.3 Four-Point Approximation

Similar analysis may be applied to derive a four-point approximation of first order derivative as follows. Let a four-point approximation be given by

$$f'(z) = \sum_{k=1}^{4} \alpha_k f(z+h_k\epsilon) + O(\epsilon^r). \tag{39}$$

Taylor expansion will be used to find $h_1, \cdots, h_4$ and $\alpha_1, \cdots, \alpha_4$ and the highest integer r such that the h_k's are distinct and each of α_k is nonzero. As in the three-point case, h_1, h_2, h_3, h_4 must solve the equation $h^4 + a_1h^3 + a_2h^2 + a_3h + a_4 = 0$, where a_1, a_2, a_3, a_4 are determined from the following equation

$$\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & \gamma_1 \\ 0 & 0 & \gamma_1 & \gamma_2 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = \begin{bmatrix} -a_4 \\ -a_3 \\ -a_2 \\ -a_1 \end{bmatrix}, \tag{40}$$

or

$$\begin{bmatrix} a_4 \\ a_3 \\ a_2 \\ a_1 \end{bmatrix} = -\begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{-\gamma_2}{\gamma_1^2} & \frac{1}{\gamma_1} \\ 0 & 0 & \frac{1}{\gamma_1} & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{bmatrix} = -\begin{bmatrix} \gamma_1 \\ 0 \\ \frac{\gamma_3}{\gamma_1} - \frac{\gamma_2^2}{\gamma_1^2} \\ \frac{\gamma_2}{\gamma_1} \end{bmatrix}. \tag{41}$$

This shows that h_1, h_2, h_3, h_4 satisfy the equation

$$h^4 - \frac{\gamma_2}{\gamma_1}h^3 + (\frac{\gamma_3}{\gamma_1} - \frac{\gamma_2^2}{\gamma_1^2})h^2 - \gamma_1 = 0. \tag{42}$$

Since $\gamma_1, \gamma_2, \gamma_3$ are arbitrary complex numbers such that $\gamma_1 \neq 0$, one may assume for convenience that $\gamma_2 = \gamma_3 = 0$, and $\gamma_1 = \gamma^4$ for some $\gamma \neq 0$, then h_1, h_2, h_3, h_4 satisfy the quartic equation $h^4 - \gamma^4 = 0$, i.e, $h = \gamma, w\gamma, w^2\gamma, w^3\gamma$, where w is a primitive 4th root of 1, (i.e., $w^4 = 1$, and $w \neq 1$, $w = \pm j$). Thus the corresponding values of $\alpha_1, \alpha_2, \alpha_3$ and α_4 are determined from the system

$$\begin{bmatrix} 1 & 1 & 1 & \\ h_1 & h_2 & h_3 & h_4 \\ h_1^2 & h_2^2 & h_3^2 & h_4^2 \\ h_1^3 & h_2^3 & h_3^3 & h_4^3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ a_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}. \tag{43}$$

This implies that

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ \gamma & j\gamma & -\gamma & -j\gamma \\ \gamma^2 & -\gamma^2 & \gamma^2 & -\gamma^2 \\ \gamma^3 & -j\gamma^3 & -\gamma^3 & j\gamma^3 \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{4\gamma} \\ \frac{-j}{4\gamma} \\ \frac{-1}{4\gamma} \\ \frac{j}{4\gamma} \end{bmatrix}. \tag{44}$$

Consequently the fourth order approximation for $f'(z)$ is given by

$$\begin{aligned} f'(z) &= \frac{1}{4\gamma} f(z+\gamma\epsilon) - \frac{j}{4\gamma} f(z+\gamma j\epsilon) - \frac{1}{4\gamma} f(z-\gamma\epsilon) \\ &+ \frac{j}{4\gamma} f(z-j\gamma\epsilon) + O(\epsilon^4). \end{aligned} \tag{45}$$

Finally, one can show that

$$\begin{aligned} f'(z) &= \frac{1}{5\gamma} f(z+\gamma\epsilon) + \frac{w^4}{5\gamma} f(z+\gamma w\epsilon) + \frac{w^3}{5\gamma} f(z+w^2\gamma\epsilon) \\ &+ \frac{w^2}{5\gamma} f(z+w^4\gamma\epsilon) + \frac{w}{5\gamma} f(z+w^4\gamma\epsilon) + O(\epsilon^5), \end{aligned} \tag{46}$$

is a five-point approximation for $f'(z)$ of order five. Here w is a primitive 5th root of 1, γ, and ϵ are nonzero numbers.

5 Multi-Point Approximation of $f''(z)$

The ideas of the previous sections can be generalized to derive multi-point approximations of $f''(z)$, i.e., to find the parameters h_k's, α_k's and r in the following

$$f''(z) = \sum_{k=1}^{N} \alpha_k f(z+h_k\epsilon) + O(\epsilon^r), \tag{47}$$

for some desired integer N. For example if w is a primitive cube root of 1 and $N = 3$, then

$$\begin{aligned} f''(z) &= \frac{2}{3\gamma^2} f(z+\gamma\epsilon) + \frac{2w}{3\gamma^2} f(z+\gamma w\epsilon) \\ &+ \frac{2w^2}{3\gamma^2} f(z+w^2\gamma\epsilon) + O(\epsilon^3), \end{aligned} \tag{48}$$

is a three-point approximation of $f''(z)$. As in the previous sections, the parametrers h_k's, and α_k's can be chosen to be real. This will be analyzed in upcoming article.

6 Multi-Point Zero-Finding Methods

One application of multi-point approximation of derivatives, is to modify existing methods such as Halley's and Ostrowski's methods as shown in the following formulas, respectively:

$$\Phi(z) = z - \frac{f(z)}{\sum_{k=1}^{N} \alpha_k f(z+h_k\epsilon) - \frac{f(z)\sum_{k=1}^{N} \alpha_k f(z+h_k\epsilon)}{2\sum_{k=1}^{N} \beta_k f(z+l_k\epsilon)}}, \tag{49}$$

and

$$\Phi(z) = z - \frac{f(z)}{\sqrt{(\sum_{k=1}^{N} \alpha_k f(z+h_k\epsilon))^2 - f(z)\sum_{k=1}^{N}\beta_k f(z+l_k\epsilon)}}. \qquad (50)$$

Here N is a positive integer greater than 2, the expressions $\sum_{k=1}^{N} \alpha_k f(z+h_k\epsilon)$, and $\sum_{k=1}^{N} \beta_k f(z+l_k\epsilon)$, are N-point approximations of $f'(z)$ and $f''(z)$ for some coefficients $h_k, l_k, \alpha_k, \beta_k, k = 1, \cdots, N$.

7 Examples

We present here two examples to show that the proposed methods can also be applied to entire functions. Let $f(z) = \sin(z)$, then

$$f(z)^2 - f(z+h)f(z-h) = \sin(z)^2 - \sin(z+h)\sin(z-h)$$
$$= \sin(z)^2 - \sin(z)^2\cos(h)^2 + \cos(z)^2\sin(h)^2 = \sin(h)^2.$$

Therefore, a third order iteration can be obtained as:

$$\Phi(z) = z - \frac{h\sin(z)}{\sqrt{\sin(h)^2}} = z - \frac{\sin(z)}{\mathrm{sinc}(h)}$$

if $h > 0$. Here $\mathrm{sinc}(h) = \frac{\sin(h)}{h}$ for $h \neq 0$ and $\mathrm{sinc}(0) = 1$.

Similarly, if $f(z) = \cos(z)$, then

$$f(z)^2 - f(z+h)f(z-h) = \cos(z)^2 - \cos(z+h)\cos(z-h)$$
$$= \cos(z)^2 - \cos(z)^2\cos(h)^2 + \sin(z)^2\sin(h)^2 = \sin(h)^2.$$

A third order iteration can be obtained as:

$$\Phi(z) = z - \frac{h\cos(z)}{\sqrt{\sin(h)^2}}.$$

It is interesting to note that the expression $f(z)^2 - f(z+h)f(z-h)$ is independent of z. This property for polynomials holds only for polynomials of degree 1 or less.

It should also be noted that as $h \to 0$ the ratio $\frac{h}{\sqrt{\sin(h)^2}} \to \pm 1$. Consequently, these iterations for the equations $\sin(z) = 0$ and $\cos(z) = 0$ redue to the square root iteration applied to these equations.

8 Conclusion

A multi-parameter derivative-free family of methods for finding simple zeros of nonlinear equations is presented. The approximation approach is carried out by approximating the logarithmic derivative of polynomials. Newton, Ostrowski, and higher order root iteration developed by the author in [16] are seen as special cases of the family. Additionally, an optimal method is given for developing multi-point approximation of first and higher order derivatives. Thus, by utilizing multi-point approximation of derivatives, one-point higher order methods can be converted to multi-point methods. Preliminary numerical computation, which will be included in the final version, indicated that the multi-point methods described in (17) and (19) converge fast when applied to polynomial equations. There are some adjustments that need to be made regarding which square or cubic root is to be chosen.

References

[1] A. S. Householder, The Numerical Treatment of a Single Nonlinear Equation, McGraw-Hill, New York, 1970.

[2] Henrici, P., Applied and computational complex analysis, Vol.I, New York: John Wiley and Sons Inc.,1974.

[3] Jarratt P., "Some Fourth Order Multipoint Iterative Methods for Solving Equations," Mathematics of Computation, Vol. 20, No. 95. (Jul., 1966), pp. 434-437.

[4] M. K. Jain, "Fifth order implicit multipoint method for solving equations," BIT Numerical Mathematics, Vol 25, Number 1, March, 1985, Pp. 250-255.

[5] A. M. Ostrowski, Solution of Equations in Euclidean and Banach Spaces, Academic, New York, 1973.

[6] J. F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, NJ, 1964.

[7] Hasan, Mohammed A., "On the Derivation of Higher Order Root-Finding Methods," American Control Conference, 2007, pp:2328-2333.

[8] E. Halley, A new, exact, and easy method of finding roots of any equations generally, and that without any previous reduction. Philos. Trans. Roy. Soc. London 18 (1694), pp. 136-145.

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

[10] Parlett B., Laguerre's method applied to the matrix eigenvalue problem, Math. Comp., 18 (1964), pp.464- 485.

[11] Foster L.V., Generalizations of Laguerre's method: Higher order methods, SIAM J.Numer.Anal., 18 (1981), pp.1004-1018.

[12] Hansen, E., Patrick, M., A family of root-finding method, Numer.Math., 27 (1977), 257-269.

[13] Bateman H., Halley's method for solving equations, Amer. Math. Monthly, 45 (1938), pp. 11-17.

[14] Jurgen Gerlach,Accelerated Convergence in Newton's Method," SIAM Review,Vol.36,No.2.(Jun.,1994),pp.272-276.

[15] P. M. Hummel; C. L. Seebeck, Jr., Generalization of Taylor's Expansion The American Mathematical Monthly, Vol. 56, No. 4. (Apr., 1949), pp. 243-247.

[16] Hasan, M.A., Root Iterations and the Computation of Minimum and Maximum Zeros of Polynomials, IEEE International Symposium on Circuits and Systems, 2005. ISCAS 2005. 23-26 May 2005, pp: 2259-2262.

[17] Dunham, W. "Cardano and the Solution of the Cubic." Ch. 6 in Journey through Genius: The Great Theorems of Mathematics. New York: Wiley, pp. 133-154, 1990.