Approximation of Polynomial Root Using a Single Input and the Corresponding Derivative Values

Bahman Kalantari
Department of Computer Science
Rutgers University, New Brunswick, NJ 08903

Abstract

A new formula for the approximation of root of polynomials with complex coefficients is presented. For each simple root there exists a neighborhood such that given any input within this neighborhood, the formula generates a convergent sequence, computed via elementary operations on the input and the corresponding derivative values. Each element of the sequence is defined in terms of the quotient of two determinants, computable via a recursive formula. Convergence is proved by deriving an explicit error estimate. For special polynomials explicit neighborhoods and error estimates are derived that depend only on the initial error. In particular, the latter applies to the approximation of root of numbers. The proof of convergence utilizes a family of iteration functions, called the Basic Family; a nontrivial determinantal generalization of Taylor’s theorem; a lower bound on determinants; Gerschgorin’s theorem and Hadamard’s inequality; as well as several new key results. The convergence results motivate a new strategy for general rootfinding, where in each iteration one approximates a root of the Taylor polynomial of a desirable degree, via the above sequence. The results also motivate the development of new sequences of iteration functions.

Keywords: Rootfinding, Polynomial Zeros, Order of Convergence.

AMS Subject Classification. 12D10, 30C15, 41A20, 65H05, 65Y99.
1 Introduction.

Polynomial rootfinding is a very interesting, deep, and applied problem, studied for many centuries. A comprehensive bibliography is given by McNamee [24]. For a recent article on some history, applications, and algorithms, see Pan [32]. Some rootfinding methods first obtain a high precision approximation to a root, then approximates other roots after deflation, see e.g., Jenkins and Traub [10]. Many such rootfinding methods make use of iteration functions, e.g. Newton’s or Laguerre’s. A method that guarantees convergence to all the roots and is a two-dimensional analogue of the bisection method is due to Weyl [34]. Modifications of this method has been used to obtain initial approximation to a root, followed by the use of Newton’s method, to give guaranteed estimate on the complexity of approximation, see e.g. Pan [31], Renegar [25].

Another rootfinding method is based on recursive factorization of the given polynomial, see e.g. Kirrinnis [21]. For many deep theoretical complexity results on polynomial rootfinding and/or the use of Newton’s method, see e.g. Smale [26, 27], Shub and Smale [28, 29], Friedman [6].

Other important topics on polynomials can be found in Bini and Pan [3], and Borwein and Erdélyi [4].

In this paper we obtain new results on the polynomial rootfinding problem: Given a single input, selected from an appropriate neighborhood of a simple root, a formula is given that generates a convergent sequence, computed via elementary operations on the input and the corresponding derivative values. Unlike the ordinary iterative methods, no other function evaluation is needed. The formula is defined in terms of determinants that can be computed directly or recursively. The convergence result motivates a new strategy for general rootfinding, where in each iteration one approximates the root of the Taylor polynomial of desirable degree, via the above mentioned sequence. The results also motivate the development of new sequences of iteration functions. Although in this paper we will not consider rootfinding for functions other than polynomials, the formula given in this paper can be used in more generality. For instance, for one possible determinantal approximation of \( \pi \), see Kalantari [14]. Other determinantal approximations to \( \pi \) will be derived in [20]. These are very different than the existing formulas for approximation of \( \pi \), such as those described in Berggren, Borwein, and Borwein [2], Pi: A Source Book.

We will next describe the new formula. Let \( f(x) \) be a polynomial of degree \( n \geq 2 \) with complex coefficients. We wish to approximate roots of \( f(x) \). For a complex number \( c = a + ib \), its modulus is \( |c| = \sqrt{a^2 + b^2} \). Given a positive number \( r > 0 \), and a complex number \( \theta \), let
\[ N_r(\theta) = \{z : |z - \theta| \leq r\} \]

Define \( D_0(x) \equiv 1 \), and for each \( m \geq 1 \), let

\[
D_m(x) = \det \begin{pmatrix}
    f'(x) & \frac{f''(x)}{2!} & \cdots & \frac{f^{(m-1)}(x)}{(m-1)!} & \frac{f^{(m)}(x)}{m!} \\
    f(x) & f'(x) & \ddots & \ddots & \ddots \\
    \vdots & \vdots & \ddots & \ddots & \ddots \\
    0 & 0 & \cdots & f(x) & f'(x)
\end{pmatrix},
\]

where \( \det \) denotes determinant. Also, given any \( m \geq 1 \), for each \( i = m+1, \ldots, n+m-1 \) define

\[
\hat{D}_{m,i}(x) = \det \begin{pmatrix}
    \frac{f''(x)}{2!} & \frac{f'''(x)}{3!} & \cdots & \frac{f^{(m)}(x)}{(m-1)!} & \frac{f^{(m+1)}(x)}{m!} \\
    f'(x) & \frac{f''(x)}{2!} & \cdots & \ddots & \ddots \\
    \vdots & \vdots & \cdots & \ddots & \ddots \\
    0 & 0 & \cdots & f'(x) & f''(x)
\end{pmatrix}.
\]

The following theorem (Theorem 1.1) is a consequence of results proved in Kalantari et. al [12], and Kalantari [14]. The main result of this paper makes use of this key theorem.

**Theorem 1.1.** For each \( m \geq 2 \), define

\[
B_m(x) \equiv x - f(x) \frac{D_{m-2}(x)}{D_{m-1}(x)}.
\]  

Let \( \theta \) be a simple root of \( f(x) \). Then,

\[
B_m(x) = \theta + \sum_{i=m}^{m+n-2} (-1)^m \hat{D}_{m-1,i}(x) (x - \theta)^i.
\]

In particular, there exists \( r > 0 \) such that given any \( x_0 \in N_r(\theta) = \{z : |z - \theta| \leq r\} \), the fixed-point iteration

\[
x_{k+1} = B_m(x_k), \quad k = 1, 2, \ldots,
\]

is well-defined, it converges to \( \theta \) having order \( m \). Specifically,

\[
\lim_{k \to \infty} \frac{(\theta - x_{k+1})^m}{(\theta - x_k)^m} = (-1)^{m-1} \hat{D}_{m-1, m}(\theta) = (-1)^{m-1} \frac{D_{m-1, m}(\theta)}{f'(\theta)^{m-1}}.
\]

In particular, Theorem 1.1 gives the order of convergence of each \( B_m(x) \), and its asymptotic error constant. The first member, \( B_2(x) \), is simply Newton’s iteration function. The next member, \( B_3(x) \), is Halley’s iteration function. The family of iteration functions \( \{B_m(x)\}_{m=2}^{\infty} \), called the Basic Family, is derived and characterized in [12]. The derivation of this family is
based on an algebraic motivation, originally developed for approximation of square and cube root of numbers in [11]. For several optimal properties of the Basic Family see [12].

Halley’s method can also be obtained by applying Newton’s method to the function $f/\sqrt{f'}$, see Bateman [1]. For more on Halley’s method see Halley [9], Traub [30], Scavo and Tho [33], and [16]. Gerlach [7], gives a generalization of Bateman’s approach, and for each $m \geq 2$, recursively defines an iteration function $G_m(x)$ of order $m$. It does not follow that $G_m(x)$ would reduce into a rational function of the input and the corresponding function and derivative values. Ford and Pennline [5], give a rational formulation of Gerlach’s $G_m(x)$, but offer no closed formula. Kalantari and Gerlach [17], establish the equivalence of these two functions making use of a key recursive formula for $D_m(x)$, derived in the present paper (see Lemma 2.1). Thus, $G_m$ enjoys all the existing results on the Basic Family, e.g., closed form, efficient computation, multipoint versions, etc.

The expansion formula given in part (2) of Theorem 1.1 is a consequence of a nontrivial determinantal generalization of Taylor’s theorem with confluent divided differences, given in [14]. In particular, this determinantal Taylor theorem gives a more general development of the Basic Family, and its multi-point version. Each member $B_m$ of the Basic Family blossoms into $m$ iteration functions, $B_m^{(1)} = B_m, B_m^{(2)}, \ldots, B_m^{(m)}$, where $B_m^{(k)}$ is a $k$-point iteration function. The order of convergence of these iteration functions is derived in [13]. In [18], we describe computational results with the first nine members of the Basic Family for random polynomials. The results indicate that Newton, secant, and Halley methods are less efficient than the next six iteration functions considered.

The main result of the present paper is the following theorem, Theorem 1.2, according to which for each simple root $\theta$ of $f(x)$, there exists a neighborhood where given an arbitrary input $x_0$ within this neighborhood, the sequence $\{B_m(x_0)\}_{m=2}^{\infty}$ converges to $\theta$. The next two theorems consider the case of special polynomials and derive specific neighborhoods, as well as error estimates.

**Theorem 1.2.** Let $f(x)$ be a polynomial with complex coefficients. Let $\theta$ be a simple root of $f(x)$. There exists $r^* \in (0, 1)$ such that given any $x_0 \in N_{r^*}(\theta)$, we have

$$\theta = \lim_{m \to \infty} B_m(x_0).$$

More precisely, let

$$w(x) = \frac{1}{f'(x)} \sqrt{\sum_{i=0}^{n} \left( \frac{|f^{(i)}(x)|}{i!} \right)^2}.$$
There exists $r^* \in (0, 1)$ such that given any $x_0 \in N_\cdot(\theta)$, for all $m \geq 2$, we have

$$|B_m(x_0) - \theta| \leq \frac{|2w(x_0)(x_0 - \theta)|^m}{1 - |x_0 - \theta|} \leq \frac{|4w(\theta)(x_0 - \theta)|^m}{1 - |x_0 - \theta|},$$

and

$$|4w(\theta)(x_0 - \theta)| < 1.$$

**Definition 1.1.** Given a positive number $\delta$, the polynomial $f(x)$ is said to satisfy the $\delta$-property over a subset $S$ of the complex numbers if given $x \in S$, we have

$$\sum_{i=0}^{n} \frac{|f^{(i)}(x)|}{i!} \leq (1 + \delta)|f'(x)|.$$

**Theorem 1.3.** Suppose that $f(x)$ satisfies the $\delta$-property over $S$, $\delta \in (0, 1)$. Then, given any $x_0 \in S$, we have

$$|B_m(x_0) - \theta| \leq \frac{1 + \delta}{1 - \delta} \left[\frac{m}{2}\right] |x_0 - \theta|^m \left(1 - |x_0 - \theta|\right).$$

In particular,

**Theorem 1.4.** Suppose that $f(x) = x^n - \alpha$, where $\alpha$ is a positive number, $n$ a natural number satisfying $\theta = \alpha^{1/n} \geq n - 1 \geq 1$. Given any $x_0 \in (\theta, \theta + \frac{1}{12}]$, we have

$$|B_m(x_0) - \theta| \leq \left(\frac{\sqrt{11}}{12}\right)^{m-1}.$$

The actual interval of convergence for the approximation of $n$-th root may be much larger than the conservative estimate of Theorem 1.4. For instance, in the approximation of square roots one can select $x_0$ to be any number between the floor and ceiling of $\sqrt{\alpha}$ and possibly even larger, see [14].

In Section 2, we derive some key properties of the matrix $D_m(x)$. In Section 3, we give the proof of Theorem 1.2, and in Section 4, proof of Theorems 1.3, and 1.4. In Section 5, we considers the algorithmic applications of Theorem 1.2, and also introduces new sequences of iteration functions called, the **Truncated Basic Family**.

### 2 Properties of $D_m(x)$.

In this section we develop several significant properties of $D_m(x)$. In fact some of these properties pertain to the determinant of any matrix whose entries are constant along each diagonal, also having all entries below the subdiagonal entries, equal to zero. The first property makes such
a matrix a Toeplitz matrix, while the second property an upper Hessneberg matrix (see Golub and Van Loan [8]).

**Lemma 2.1.** Let \(a_0, \ldots, a_n\) be arbitrary complex numbers, \(n \geq 1\). Define \(a_i = 0\), for all \(i > n\). Given a natural number \(m\), define

\[
d_m = \det \begin{pmatrix} a_1 & a_2 & \cdots & a_{m-1} & a_m \\ a_0 & a_1 & & & a_{m-1} \\ & a_0 & \ddots & & \vdots \\ & & \ddots & \ddots & a_2 \\ & & & a_0 & a_1 \end{pmatrix}.
\]

Then,

\[
d_m = \sum_{i=1}^{n} (-1)^{i-1} a_0^{i-1} a_i d_{m-i},
\]

where \(a_0 \equiv 1\), \(a_0 \equiv 1\), and \(d_i \equiv 0\), if \(i < 0\). In particular, for all \(m \geq 1\) we have,

\[
D_m = \sum_{i=1}^{n} (-1)^{i-1} \frac{i!}{i!} f(i) D_{m-i},
\]

where \(D_j = 0\), for \(j < 0\).

**Proof.** Given any pair of natural numbers \(m\) and \(k\), define

\[
d_{m}^{(k)} = \det \begin{pmatrix} a_k & a_{k+1} & \cdots & a_{m+k-2} & a_{m+k-1} \\ a_0 & a_1 & & & a_{m-1} \\ & a_0 & \ddots & & \vdots \\ & & \ddots & \ddots & a_2 \\ & & & a_0 & a_1 \end{pmatrix}.
\]

Note that when \(k = 1\), \(d_{m}^{(k)} = d_m\). By expanding along the first column, the determinant of the matrix corresponding to \(d_{m}^{(k)}\), we get

\[
d_{m}^{(k)} = a_k d_{m-1} - a_0 d_{m-1}^{(k+1)}.
\]

By expanding along the first column the determinant of the matrix corresponding to \(d_m\), we get

\[
d_m = a_1 d_{m-1} - a_0 d_{m-1}^{(2)}.
\]

Now by using in the above the previously given formula for various values of \(m\) and \(k\) repeatedly, the proof of the lemma follows. \(\Box\)

The following is a key result.
Theorem 2.1. Let \( a_0, \ldots, a_n \) be arbitrary complex numbers, \( n \geq 1 \). Define \( a_i = 0 \), for all \( i > n \). For each \( i \) define

\[
b_i = a_0^{-1} a_i,
\]

where \( a_0^0 \equiv 1 \). Given a natural number \( m \), let \( d_m \) be as in Lemma 2.1, and define

\[
c_m = \det \begin{pmatrix} b_1 & b_2 & \ldots & b_{m-1} & b_m \\ b_0 & b_1 & \ldots & b_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & b_0 & b_1 \end{pmatrix}.
\]

Then, \( d_m = c_m \). In particular,

\[
D_m(x) = \det \begin{pmatrix} f' & \sqrt{T^2} f'' & \ldots & \sqrt{T} f^{(m-1)} & f^{(m)} \\ \sqrt{T} & f' & \ldots & \sqrt{T} f^{(m-1)} & f^{(m)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & \sqrt{T} & \ldots & \sqrt{T} f'' & f' \\ 0 & 0 & \ldots & \sqrt{T} & f' \end{pmatrix}.
\]

Proof. From Lemma 2.1, \( c_m \) satisfies the linear recurrence relation

\[
c_m = \sum_{i=1}^{n} (-1)^{i-1} b_0^{i-1} b_i c_{m-i},
\]

where, \( b_0^0 = 1 \), \( c_0 \equiv 1 \), and \( c_i \equiv 0 \), if \( i < 0 \). But from the definition of \( b_i \) we have

\[
c_m = \sum_{i=1}^{n} (-1)^{i-1} a_0^{-1} a_i^{-1} a_0^{i-1} a_i c_{m-i} = \sum_{i=1}^{n} (-1)^{i-1} a_0^{-1} a_i c_{m-i}.
\]

Thus, \( c_m = d_m \), since they satisfy the same recurrence relation and the same initial conditions. \( \square \)

Theorem 2.2. Let \( a_0, \ldots, a_n \) be arbitrary complex numbers, where \( a_1 \neq 0 \). Define \( a_i = 0 \), for all \( i > n \). For each natural number \( m \), let \( d_m \) be as in Lemma 2.1. There exists \( \epsilon > 0 \), such that if \( |a_0| \leq \epsilon \), then for all \( m \geq 1 \), we have

\[
|d_m| \geq \left( \frac{|a_1|}{2} \right)^m.
\]

Proof. The equivalence of \( d_m \) and \( c_m \) will be used to derive the claimed lower bound. From the definition of \( b_i \), we have \( b_1 = a_1 \). For \( i \neq 1 \), \( |b_i| \) can be made arbitrarily small, assuming that \( |a_0| \) is sufficiently small. Thus, there exists \( \epsilon > 0 \) such that if \( |a_0| \leq \epsilon \), then

\[
|a_1| - \sum_{i=0, i \neq 1}^{m} |b_i| \geq \frac{|a_1|}{2} > 0.
\]
In particular, if \( C_m \) is the matrix corresponding to \( c_m \), and \( |a_0| \leq \epsilon \), then \( C_m \) is diagonally dominant. Hence, from Gerschgorin’s theorem it follows that if \( \lambda_1, \ldots, \lambda_m \) are the eigenvalues of \( C_m \), then \( |\lambda_i| \geq |a_1|/2 \). Consequently, when \( |a_0| \leq \epsilon \), we have

\[
|d_m| = |\text{det}(C_m)| = |\lambda_1 \cdots \lambda_m| \geq \frac{|a_1|^m}{2^m}. \quad \square
\]

**Theorem 2.3.** (Kalantari \([15]\)) Let \( A \) be an \( m \times m \) complex matrix. Assume we are given positive numbers \( l \) and \( u \) such that if \( \lambda \) is an eigenvalue of \( A \), then \( l \leq |\lambda| \leq u \). Let

\[
\kappa = \kappa(l, u) \equiv \left\lceil \frac{mu - |\text{trace}(A)|}{u - l} \right\rceil.
\]

If \( |\text{trace}(A)| \geq ml \), then

\[
|\text{det}(A)| \geq l^\kappa u^{m-\kappa}. \quad \square
\]

**Theorem 2.4.** Let \( a_0, \ldots, a_n \) be arbitrary complex numbers. Define \( a_i = 0 \), for all \( i > n \). For each natural number \( m \) let \( d_m \) be as in Lemma 2.1. Let \( \Delta_m \) be the matrix corresponding to \( d_m \). Suppose that \( \Delta_m \) is diagonally dominant. Let

\[
\delta = \frac{1}{|a_1|} \sum_{i=0, i\neq 1}^m |a_i|.
\]

Then,

\[
|d_m| \geq (1 - \delta)^\left\lceil \frac{m}{2} \right\rceil (1 + \delta)^\left\lfloor \frac{m}{2} \right\rfloor |a_1|^m.
\]

**Proof.** From Gerschgorin’s theorem it follows that if \( \lambda \) is an eigenvalue of \( \Delta_m \), we have

\[
l = (1 - \delta)|a_1| \leq \lambda \leq (1 + \delta)|a_1| = u.
\]

Since \( |\text{trace}(\Delta_m)| = m|a_1| \geq (1 - \delta)|a_1| \), Theorem 2.3 applies to \( \Delta_m \), where

\[
\kappa = \left\lceil \frac{mu - |\text{trace}(\Delta_m)|}{u - l} \right\rceil = \left\lceil \frac{m(1 + \delta)|a_1| - m|a_1|}{(1 + \delta)|a_1| - (1 - \delta)|a_1|} \right\rceil = \left\lceil \frac{m}{2} \right\rceil.
\]

Thus, from Theorem 2.3 we have,

\[
|d_m| \geq \left( (1 - \delta)|a_1| \right)^\left\lceil \frac{m}{2} \right\rceil \left( 1 + \delta |a_1| \right)^\frac{m}{2}.
\]

which coincides with the claimed lower bound. \( \square \)
3 Proof of Theorem 1.2.

From Theorem 2.2, there exists \( \epsilon > 0 \) such that if \( |f(x)| \leq \epsilon \), then for all \( m \geq 1 \), we have

\[
|D_m(x)| \geq \left( \frac{|f'(x)|}{2} \right)^m.
\]

But since \( \theta \) is a simple root of \( f(x) \), there exists \( r_1 \in (0, 1) \) such that if \( x \in N_{r_1}(\theta) \), then \( |f(x)| \leq \epsilon \), and \( f'(x) \neq 0 \). Define

\[
h(x) = \sqrt{\sum_{i=0}^{n} \left( \frac{|f^{(i)}(x)|}{i!} \right)^2}.
\]

Consider the matrix corresponding to \( \hat{D}_{m,i}(x) \) (see Section 1). The norm of each of its columns is bounded above by \( h(x) \). Given a matrix \( A = (a_{ij}) \), from Hadamard’s inequality (see e.g. Marcus and Minc [23]) we have

\[
|\det(A)| \leq \left( \prod_{i=1}^{n} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{1/2}.
\]

Thus, we have

\[
|\hat{D}_{m,i}(x)| \leq h(x)^m.
\]

Using the upper bound on \( |\hat{D}_{m,i}(x)| \), the lower bound on \( |D_m(x)| \), part (2) of Theorem 1.1, and the geometric progression formula, we conclude that for each \( x \in N_{r_1}(\theta) \), we have

\[
|B_m(x) - \theta| \leq \frac{2h(x)}{|f'(x)|} |x - \theta|^m \sum_{i=0}^{\infty} |x - \theta|^i \leq \frac{2h(x)}{|f'(x)|} |x - \theta|^m \frac{1}{1 - |x - \theta|}.
\]

Let \( w(x) = h(x)/f'(x) \). From continuity, there exist \( r_2 > 0 \) such that if \( x \in N_{r_2}(\theta) \), then \( |w(x)| \leq |2w(\theta)| \), and \( |2w(\theta)(x - \theta)| < 1 \). Define \( r^* = \min\{r_1, r_2\} \). From the above bound it follows that for any \( x_0 \in N_{r^*}(\theta) \), \( B_m(x_0) \) converges to \( \theta \). \( \Box \)

4 Proof of Theorems 1.3 and 1.4.

To prove Theorem 1.3, first we give a lemma.

**Lemma 4.1.** Suppose that \( f(x) \) satisfies the \( \delta \) property over a set \( S, \delta \in (0, 1) \). Then, for any \( x_0 \in S \), we have

\[
|w(x_0)| = \frac{h(x_0)}{|f'(x_0)|} = \frac{1}{|f'(x_0)|} \sqrt{\sum_{i=0}^{n} \left( \frac{|f^{(i)}(x_0)|}{i!} \right)^2} \leq (1 + \delta),
\]
\[ |\hat{D}_{m,i}(x_0)| \leq \left( (1+\delta)|f'(x_0)| \right)^m, \]

and

\[ |D_m(x_0)| \geq (1-\delta)^{\frac{m}{2}}(1+\delta)^{\frac{m}{2}}|f'(x_0)|^m. \]

**Proof.** Clearly,

\[ h(x_0) \leq \sum_{i=0}^{n} \frac{|f^{(i)}(x_0)|}{i!} \leq (1+\delta)|f'(x_0)|. \]

The above implies the first inequality of the lemma. Recall that from Hadamard’s inequality \(|\hat{D}_{m,i}(x_0)| \leq h(x_0)^m\). Hence, the second inequality of the lemma follows. The third inequality follows from Theorem 2.4. \(\square\).

From the above lemma, together with Theorem 1.1, and analogous to the proof of Theorem 1.2, for any \(x \in S\) we obtain the bound

\[ |B_m(x) - \theta| \leq \left| \frac{h(x)}{D_m(x)} \right|^m \frac{|x-\theta|^m}{1-|x-\theta|} \leq \frac{(1+\delta)^m|x-\theta|^m}{(1-\delta)^{\frac{m}{2}}(1+\delta)^{\frac{m}{2}}}. \]

Clearly, the above upper bound is equivalent to the claimed upper bound in Theorem 1.3.

To prove Theorem 1.4, we first prove a lemma.

**Lemma 4.2.** Let \(\alpha\) be a positive number, and \(n\) a natural number satisfying \(\theta = \alpha^{1/n} \geq n - 1 \geq 1\), and \(\delta = \frac{5}{6}\). Then, \(f(x) = x^n - \alpha\) satisfies the \(\delta\)-property over the interval \((\theta, \theta + \frac{1}{12}]\).

**Proof.** Let \(x_0 \in (\theta, \theta + \frac{1}{12}]\). For each \(i = 2, \ldots, n\), we have

\[ f^{(i)}(x_0) = n(n-1) \cdots (n-i+1)x_0^{n-i} \leq nx_0^{n-1} = f'(x_0). \]

Also,

\[ f(x_0) = x^n - \theta^n = (x_0 - \theta)(\sum_{i=0}^{n-1} x_0^{n-i} \theta^i) \leq (x_0 - \theta)nx_0^{n-1} = (x_0 - \theta)f'(x_0) \leq \frac{f'(x_0)}{12}. \]

Thus,

\[ \frac{1}{f'(x_0)} \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!} \leq \left( 1 + \frac{1}{12} + \frac{1}{6} \right). \]

But

\[ \left( 1 + \frac{1}{12} + \frac{1}{6} \right) = 1 + \frac{1}{12} + \frac{2}{12} + \frac{1}{6} + \frac{1}{24} + \frac{n}{24} i! \leq 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} \sum_{i=0}^{\infty} \frac{1}{2^i} = 1 + \frac{5}{6} \square \]

Now to prove Theorem 1.4, from the above lemma, the fact that \(x_0 - \theta \leq \frac{1}{12}\), and Theorem 1.3, we have

\[ |B_m(x_0) - \theta| \leq \left( 1 + \frac{5}{6} \right) \left( \frac{\frac{m}{2}}{1-\frac{5}{6}} \right)^m \leq \frac{11^{\frac{m}{2}}}{12^{m-1}} - \frac{11^{m-1}}{12^{m-1}} = \left( \frac{\sqrt{11}}{12} \right)^{m-1}. \]
5 Applications.

The first application of the convergence result of Theorem 1.2 is a new method for polynomial rootfinding. Given a good approximation $x_0$ to a root of the given polynomial, we make use of the sequence $\{B_m(x_0)\}_{m=2}^{\infty}$ to obtain better approximations. Theorem 1.2 implies that if $x_0$ is a good approximation to a given simple root $\theta$, the sequence $\{B_m(x_0)\}_{m=2}^{\infty}$ converges to $\theta$ and there is no need to do additional function/derivative evaluations. Even if the sequence is not convergent to $\theta$, by evaluating sufficient number of elements, we postpone function evaluations until we have drawn sufficient information from the current input $x_0$. We will formalize this in the form an algorithm (Rootfinding(n) below). Once $B_m(x_0)$ is evaluated, the evaluation of $B_{m+1}(x_0)$ can be achieved in $O(mn)$ arithmetic operations (and possibly even faster). This follows from the recursive formula of Lemma 2.1. In particular, for fixed polynomial degree, the evaluation of $B_{m+1}(x_0)$ can be established in $O(m)$ time. The entries of $B_m(x_0)$ can be evaluated efficiently since given $f(x)$, a polynomial of degree $n$, and an input $x_0$, the evaluation of $f^{(j)}(x_0)/j!$, $j = 0, \ldots, n$, called the normalized derivatives, can be established in $O(n \log^2 n)$ arithmetic operations, see Kung [22]. The recursive formula of Lemma 2.1 is also attractive from the point of view of parallel evaluation of successive elements of the sequence $\{B_m(x_0)\}_{m=2}^{\infty}$.

Aside from polynomial rootfinding, the convergence result also motivates a new strategy for the approximation of roots of more general functions. The strategy is as follows: given the current input $x_0$, one approximates a root of the corresponding $n$-th degree Taylor polynomial, for an appropriately selected degree $n$, and via the above mentioned sequence. The new approximation replaces the old one, and the process is repeated. These results also motivate the development of new sequences of iteration functions. We will first describe these new iteration functions. For each natural number $t$, it is possible to define a family of iteration function $\{B_{m,t}(x)\}_{m=t+1}^{\infty}$, where $B_{m,t}(x)$ is obtained from $B_m(x)$ by setting all derivatives higher than the $t$-th derivative equal to zero. We shall refer to $\{B_{m,t}(x)\}_{m=t+1}^{\infty}$ as the Truncated Basic Family of order $t$. It can be shown that the order of convergence of each member of this family is $t+1$. The proof of this fact will be given in [19] where we will also consider its multipoint version. The Truncated Basic Family of order 2, analyzed in [16], gives an infinite family of cubic rate methods, called Halley Family. From Theorem 1.1 it follows that for polynomials of degree at most $t$, the order of convergence of $B_{m,t}(x)$ is $m$. This is an interesting property of the Truncated Basic Family. Unlike the Halley Family which is an infinite family, the family $\{B_{m,1}(x)\}_{m=2}^{\infty}$ consists of a single element, namely Newton’s iteration function.
Now consider $f(x)$, say a real-valued function having $n$ continuously differentiable derivatives in a neighborhood of a root, and the following algorithm:

**Rootfinding**($n$):

**Step 1.** Given $x_0$, compute $P_n(x)$, the $n$-th degree Taylor Polynomial of $f$ at $x_0$.

**Step 2.** Approximate a root of $P_n(x)$ using the sequence $\{B_{m,n}(x_0)\}_{m=n+1}^{\infty}$:

- **Strategy I.** Let $x_1 = B_{n+1,n}(x_0)$.
- **Strategy II.** Fix $\epsilon > 0$. Let $x_1 = B_{m,n}(x_0)$, where $m$ is the least $m \geq n + 2$ such that $|B_{m,n}(x_0) - B_{m-1,n}(x_0)| \leq \epsilon$.

**Step 3.** Replace $x_0$ with $x_1$. Go to Step 1.

When $n = 1$, the above algorithm is simply Newton’s method. For $n \geq 2$, the approximation of the zero of the Taylor polynomial in Step 2 is via a single approximation $B_{n+1,n}(x_0)$, according to Strategy I, or the sequence of approximations $\{B_{m,n}(x_0)\}_{m=n+1}^{\infty}$, according to Strategy II. The justification in utilizing the sequence $\{B_{m,n}(x_0)\}_{m=n+1}^{\infty}$ lies in the fact that according to Theorem 1.2 when $x_0$ is a reasonably good approximation to a root of $P_n(x)$, then the sequence converges to that root. Note that the sequence $\{B_{m,n}(x_0)\}_{m=n+1}^{\infty}$ correspond to the values of the Basic Family as written with respect to the Taylor polynomial $P_n(x)$. From the recursive property of $D_m(x)$, see Lemma 2.1, once $B_{n+1,n}(x_0)$ is evaluated, the subsequent element of the sequence can be computed efficiently, and so on. Strategy II of Step 2 allows the possibility of refining the initial approximation, $B_{n+1,n}(x_0)$, by performing only elementary operations on $x_0$, and the computed values, $f^{(j)}(x_0)/j!$, $j = 0, \ldots, n$.

**References**


