Symmetric functions and root-finding algorithms

Yi Jin *, Bahman Kalantari

Department of Computer Science, Rutgers University (New Brunswick), Hill Center–Busch Campus,
110 Frelinghuysen Rd., Piscataway, NJ, 08854, USA

Received 27 February 2004; accepted 10 May 2004
Available online 5 November 2004

Abstract

We study a fundamental family of root-finding iteration functions in the context of symmetric functions. This family, which we refer to as the Basic Family, goes back to Schröder’s 1870 paper, and admits numerous different representations. In one representation, it is known as König’s family. A purely algebraic derivation by Kalantari et al. leads to the discovery of many minimality and uniqueness properties of this family. Our new perspective reveals a symmetric algebraic structure of the Basic Family, which gives rise to simple combinatorial proofs of many important properties of this family and two of its variants. The first variant maintains high order of convergence for multiple roots. The second variant, called the Truncated Basic Family, is an infinite family of $m$th order methods for every $m \geq 3$, using only the first $m - 1$ derivatives. Our result extends Kalantari’s analysis of Halley’s family, the special case of the Truncated Basic Family where $m = 3$. Finally, we give a recipe for constructing new high order root-finding algorithms, and use it to derive an interesting family of iteration functions that have higher orders of convergence for multiple roots than for simple roots.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Root-finding; Iteration functions; Schröder’s algorithms; König’s family; Symmetric functions; Schur functions

* Corresponding author.
E-mail addresses: yjin@paul.rutgers.edu (Y. Jin), kalantar@cs.rutgers.edu (B. Kalantari).

0196-8858/$ – see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.aam.2004.05.005
1. Introduction

For a general polynomial equation of degree higher than 4, there is no algebraic solution. In this case, rational iteration functions such as Newton’s method are widely used to approximate the roots. It is well known that Newton’s method has local quadratic convergence rate to simple roots. In fact, there are arbitrarily high order methods for both simple and multiple roots.

In this paper, we shall study a fundamental family of such high order methods. This family, which we refer to as the Basic Family, goes back to Schröder’s 1870 paper [17]. In that remarkable paper, Schröder gave a general theory on rational root-finding algorithms and derived many high order methods which include this fundamental family.

The Basic Family has been derived through many different approaches and admits numerous different representations. In one representation, it is known as Königs’s family (see [1,21]). In [3], Gerlach devised a scheme to accelerate Newton’s method, which turns out to be another representation of the Basic Family (see [2,9]). In [6], Kalantari et al. derived and characterized the Basic Family in a purely algebraic manner. The derivation reveals many new and interesting minimality and uniqueness properties of this family. For some related works in this direction, see [8,10,11]. For a sample of works on general root-finding methods, see [4,15,18,20].

A simple representation of the Basic Family is given in [6,11]. Let \( p(x) \) be a polynomial of degree \( n \). Define \( D_0(x) = 1 \), and for each integer \( m \geq 1 \), let

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

where \( p^{(i)}(x) \) is the \( i \)th derivative of \( p(x) \). Then, for each integer \( m \geq 2 \), the rational function

\[
B_m(x) = x - p(x) \frac{D_{m-2}(x)}{D_{m-1}(x)}
\]

defines an iterative algorithm: \( x_k = B_m(x_{k-1}) \), \( k = 1, 2, \ldots \), and \( \{x_k\}_{k=0}^{\infty} \) converges to a root of \( p \), given an appropriate initial point \( x_0 \). Moreover, its order of convergence is \( m \) for a simple root.

The function \( D_m(x) \) also has a closed formula:

\[
D_m(x) = \det \begin{bmatrix}
p(x) & p'(x) & \cdots & p^{(m-1)}(x) \frac{1}{(m-1)!} & \frac{p^{(m)}(x)}{m!} \\
p(x) & p'(x) & \cdots & \frac{p^{(m-1)}(x)}{(m-1)!} & \frac{p^{(m)}(x)}{m!} \\
0 & p(x) & \cdots & \frac{p^{(m-1)}(x)}{(m-1)!} & \frac{p^{(m)}(x)}{m!} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & \cdots & p(x) & p'(x)
\end{bmatrix},
\]
where $\det A$ denotes the determinant of matrix $A$.

The first three members of the Basic Family are

$$B_2(x) = x - \frac{p(x)}{p'(x)}, \quad (1.3)$$

which is the well-known Newton’s method, and

$$B_3(x) = x - \frac{2p(x) p'(x)}{2p'(x)^2 - p(x)p''(x)}, \quad (1.4)$$

which is Halley’s method, and the 4th order method

$$B_4(x) = x - \frac{6p(x) p'(x)^2 - 3p(x)^2 p''(x)}{p(x)^2 p'''(x) + 6p'(x)^3 - 6p(x) p'(x) p''(x)}. \quad (1.5)$$

More generally, the Basic Family can be defined for any analytic function (see [6,8]). In this paper, we restrict our study to the Basic Family corresponding to a polynomial. This allows us to make a connection between Basic Family and symmetric functions.

Our exposition is organized as follows. In Section 2 we give a brief review of symmetric function theory. In Section 3 we reveal symmetric algebraic structures of the Basic Family and two of its variants, and give combinatorial proofs of their convergence properties and some other results. In Section 3.1, we consider the Basic Family and reveal a symmetric structure different from the more restrictive one given by Schröder [17]. This symmetric structure gives rise to elementary derivation of the orders of convergence and asymptotic error constants for both simple and multiple roots, and serves as a canonical form for establishing the equivalence of different representations of the Basic Family. Finally, we make a connection between Schröder’s symmetric characterization and ours through Schur functions. In Section 3.2, we consider a variant of Basic Family which maintains high order of convergence for multiple roots. In Section 3.3, we consider the Truncated Basic Family, an infinite family of $m$th order methods for every $m \geq 3$, using only the first $m-1$ derivatives. Our result extends Kalantari’s analysis of Halley’s family (in [7]), the special case of the Truncated Basic Family where $m = 3$. Our approach demonstrates the power of algebraic combinatorial techniques in attacking challenging problems in the seemingly remote area of polynomial root-finding. In Section 4 we give a recipe for constructing new families of high order root-finding methods, and use it to derive a family of algorithms which, contrary to typical methods, have higher orders of convergence for multiple roots than for simple roots.

2. A brief review of symmetric function theory

In this section, we transcribe some relevant ingredients in the theory of symmetric functions from the definitive treatment by MacDonald [13]. For a quick introduction to symmetric functions, see Sagan [16, Chapter 4].
Consider the ring \( \mathbb{Z}[x_1, \ldots, x_n] \) of polynomials in \( n \) independent variables \( x_1, \ldots, x_n \) with integer coefficients. A polynomial in this ring is symmetric if it is invariant under permutations of \( x_1, x_2, \ldots, x_n \). The symmetric polynomials in \( \mathbb{Z}[x_1, \ldots, x_n] \) form a graded subring \( A_n \), that is, we have

\[
A_n = \bigoplus_{k \geq 0} A_n^k,
\]

where \( A_n^k \) consists of the homogeneous symmetric polynomials of degree \( k \), together with the zero polynomial, and \( \bigoplus \) denotes the direct sum.

**Definition 2.1 (Partition).** A partition is any (finite or infinite) sequence

\[
\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_r, \ldots)
\]

(2.1)

of non-negative integers in decreasing order:

\[
\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r \geq \cdots
\]

and containing only finitely many non-zero terms.

The non-zero \( \lambda_i \) in (2.1) are called the parts of \( \lambda \). The number of parts is the length of \( \lambda \), denoted by \( l(\lambda) \); and the sum of the parts is the weight of \( \lambda \), denoted by \( |\lambda| \):

\[
|\lambda| = \lambda_1 + \lambda_2 + \cdots
\]

**Definition 2.2 (Monomial symmetric functions).** Let \( \lambda \) be any partition of length \( \leq n \). The polynomial

\[
m_{\lambda}(x_1, \ldots, x_n) = \sum x_1^{\alpha_1} \cdots x_n^{\alpha_n}
\]

summed over all distinct permutations \( (\alpha_1, \ldots, \alpha_n) \) of \( \lambda = (\lambda_1, \ldots, \lambda_n) \), is clearly symmetric, and the \( m_{\lambda} \), as \( \lambda \) runs through all partitions of length \( \leq n \), form a \( \mathbb{Z} \)-basis of \( A_n \). Hence the \( m_{\lambda} \) such that \( l(\lambda) \leq n \) and \( |\lambda| = k \) form a \( \mathbb{Z} \)-basis of \( A_n^k \).

Through specialization (mapping some of the \( x_i \) to 0), one finds that for most purposes, the number of variables in a symmetric polynomial is irrelevant, provided that it is large enough, and it is often more convenient to work with symmetric “polynomials” in infinitely many variables. These symmetric pseudo-polynomials are actually formal infinite sums of monomials. They are, however, finite integral linear combinations of monomial symmetric functions \( m_{\lambda} \), and are also known as symmetric functions.
Definition 2.3 (Elementary symmetric functions). For each integer \( r \geq 0 \), the \( r \)th elementary symmetric function \( e_r \) is the sum of all products of \( r \) distinct variables \( x_i \), that is, \( e_0 = 1 \) and

\[
e_r = \sum_{i_1 < i_2 < \cdots < i_r} x_{i_1} x_{i_2} \cdots x_{i_r}, \quad \text{for } r \geq 1.
\]

It is easy to see that \( e_r = m_{1^r} \), where \( 1^r \) denotes the partition \( (1, \ldots, 1) \).

The generating function for \( e_r \) is

\[
E(t) = \sum_{r \geq 0} e_r t^r = \prod_{i \geq 1} (1 + x_i t), \quad \text{(2.2)}
\]
as one sees by multiplying out the product on the right.

Definition 2.4 (Complete symmetric functions). For each integer \( r \geq 0 \), the \( r \)th complete symmetric function \( h_r \) is the sum of all monomials of total degree (sum of exponents of all its variables) \( r \) in the variables \( x_1, x_2, \ldots \), that is,

\[
h_r = \sum_{|\lambda| = r} m_\lambda. \quad \text{(2.3)}
\]

In particular, \( h_0 = 1 \) and \( h_1 = e_1 \).

The generating function for \( h_r \) is

\[
H(t) = \sum_{r \geq 0} h_r t^r = \prod_{i \geq 1} (1 - x_i t)^{-1}, \quad \text{(2.4)}
\]
which is easy to see once we replace \((1 - x_i t)^{-1}\) by their power series expansions

\[
\sum_{k \geq 0} x_i^k t^k.
\]

From (2.2) and (2.4) we have

\[
H(t) E(-t) = 1,
\]
which implies

\[
\sum_{r=0}^{n} (-1)^r e_r h_{n-r} = 0, \quad \forall n \geq 1. \quad \text{(2.5)}
\]
**Definition 2.5 (Power sum).** For each integer \( r \geq 1 \), the \( r \)th power sum is

\[
q_r = \sum_{i \geq 1} x_i^r = m(r).
\]

The generating function for \( q_r \) is

\[
Q(t) = \sum_{r \geq 1} q_r t^{r-1} = \sum_{i \geq 1} \sum_{r \geq 1} x_i^r t^{r-1} = \sum_{i \geq 1} \frac{x_i}{1 - x_i t}.
\] (2.6)

Thus,

\[
Q(-t) = \sum_{i \geq 1} \frac{x_i}{1 + x_i t} = \frac{d}{dt} \log E(t) = \frac{E'(t)}{E(t)}
\]

which gives rise to the following identities known as Newton’s formulas:

\[
n e_n = \sum_{r=1}^{n} (-1)^{r-1} q_r e_{n-r}, \quad \forall n \geq 1.
\] (2.7)

If the number of variables is finite, say \( n \), then we write the specialized version of \( e_r, h_r, \) and \( q_r \) as \( e_r(x_1, \ldots, x_n) \), \( h_r(x_1, \ldots, x_n) \), and \( q_r(x_1, \ldots, x_n) \), respectively. It is easy to verify that both (2.5) and (2.7) hold for finite number of variables as well.

**3. Main results**

Let \( p(x) \) be a polynomial of degree \( n \) with complex coefficients, and \( \theta_1, \ldots, \theta_n \) be its complex roots. Then,

\[
p(x) = c \prod_{i=1}^{n} (x - \theta_i),
\]

where \( c \) is the leading coefficient of \( p(x) \).

Define \( r_j = 1/(x - \theta_j) \), \( j = 1, \ldots, n \), then we have

**Lemma 3.1.**

\[
p^{(i)}(x) = i! \ p(x) e_i(r_1, \ldots, r_n),
\]

where \( e_i \) is the \( i \)th elementary symmetric function.
Proof. Applying the product rule of differentiation repeatedly, we get
\[
p^{(i)}(x) = \sum_{1 \leq j_1 < \cdots < j_i \leq n} \frac{i!}{\prod_{k=1}^{i} (x - \theta_{j_k})} \prod_{k=1}^{i} r_{j_k} = i! \ p(x) \sum_{1 \leq j_1 < \cdots < j_i \leq n} \prod_{k=1}^{i} r_{j_k}.
\]

3.1. The Basic Family

We first derive a symmetric algebraic structure for the Basic Family, then use it to establish the equivalence of the Basic Family and König’s family, and prove the convergence property of the Basic Family. Via Schur functions, this symmetric structure can be transformed into Schröder’s symmetric characterization of the Basic Family for \( p \) with simple roots.

Theorem 3.2. For each integer \( m \geq 0 \),
\[
D_m(x) = p(x)^m h_m(r_1, \ldots, r_n),
\]
where \( h_m \) is the \( m \)th complete symmetric function.

Proof. We prove this theorem by induction on \( m \). Trivially, (3.1) holds for \( m = 0 \). Now assume that (3.1) holds for \( m \leq k, k \in \mathbb{Z}, k \geq 0 \). Then for \( m = k + 1 \), we have
\[
D_{k+1}(x) = \sum_{i=1}^{k+1} (-1)^{i-1} p(x)^{i-1} \frac{p^{(i)}(x)}{i!} D_{k+1-i}(x)
= \sum_{i=1}^{k+1} (-1)^{i-1} p(x)^{i-1} p(x) e_i(r_1, \ldots, r_n) p(x) h_{k+1-i}(r_1, \ldots, r_n)
= p(x)^{k+1} \sum_{i=1}^{k+1} (-1)^{i-1} e_i(r_1, \ldots, r_n) h_{k+1-i}(r_1, \ldots, r_n).
\]

From (2.5), we have
\[
\sum_{i=0}^{k+1} (-1)^i e_i h_{k+1-i} = 0.
\]

Moving the first summand on the left to the right and dividing both sides by \((-1)\), we get
\[
\sum_{i=1}^{k+1} (-1)^{i-1} e_i h_{k+1-i} = e_0 h_{k+1} = h_{k+1}.
\]
Thus,

\[ D_{k+1}(x) = p(x)^{k+1}h_{k+1}(r_1, \ldots, r_n). \]

So (3.1) holds for \( m = k + 1 \). And by induction, it holds for all \( m > 0 \). \( \Box \)

Substituting (3.1) into (1.2), we get

\[ B_m(x) = x - \frac{h_{m-2}(r_1, \ldots, r_n)}{h_{m-1}(r_1, \ldots, r_n)}. \]

(3.3)

The above identity reveals a symmetric structure of the Basic Family, and serves as a canonical form that connects different representations of this fundamental family of root-finding algorithms.

**Corollary 3.3.** Königs family is equivalent to the Basic Family, that is, \( B_m(x) \) is the same as the Königs method of order \( m \), defined by the formula

\[ K_m(x) = x + (m - 1) \frac{[1/p(x)]^{(m-2)}}{[1/p(x)]^{(m-1)}}, \]

where \([1/p(x)]^{(k)}\) is the \( k \)th derivative of \( 1/p(x) \).

**Proof.** Observe that

\[ \frac{1}{p(x)} = \frac{1}{c} \prod_{i=1}^{n} r_i, \]

and applying the product rule of differentiation repeatedly, we get

\[
\left[ \frac{1}{p(x)} \right]^{(k)} = \frac{1}{c} \sum_{d_1, \ldots, d_n \geq 0 \atop d_1 + \cdots + d_n = k} \frac{k!}{d_1! \cdots d_n!} \prod_{i=1}^{n} r_i^{(d_i)} \\
= \frac{1}{c} \sum_{d_1, \ldots, d_n \geq 0 \atop d_1 + \cdots + d_n = k} \frac{k!}{d_1! \cdots d_n!} \prod_{i=1}^{n} (-1)^{d_i} d_i! r_i^{d_i+1} \\
= \frac{(-1)^k k!}{p(x)} \sum_{d_1, \ldots, d_n \geq 0 \atop d_1 + \cdots + d_n = k} \prod_{i=1}^{n} r_i^{d_i} = \frac{(-1)^k k!}{p(x)} h_k(r_1, \ldots, r_n).
\]

Thus,

\[ K_m(x) = x + (m - 1) \frac{[1/p(x)]^{(m-2)}}{[1/p(x)]^{(m-1)}} = x - \frac{h_{m-2}(r_1, \ldots, r_n)}{h_{m-1}(r_1, \ldots, r_n)} = B_m(x). \]
The following result on the convergence rate of $B_m$ also appears in [1,6]. Here we present a new, combinatorial proof.

**Corollary 3.4.** For a simple root $\theta_1$, $B_m(x)$ has an order of convergence $m$, and

$$
\lim_{x \to \theta_1} \frac{B_m(x) - \theta_1}{(x - \theta_1)^m} = h_{m-1}(\theta_1 - \theta_2)^{-1}, \ldots, (\theta_1 - \theta_n)^{-1}).
$$

For a multiple root $\theta_1$, $B_m(x)$ converges linearly, and

$$
\lim_{x \to \theta_1} \frac{B_m(x) - \theta_1}{x - \theta_1} = \frac{s - 1}{s + m - 2},
$$

where $s \geq 2$ is the multiplicity of root $\theta_1$.

**Proof.** If $\theta_1$ is simple, then $\theta_i \neq \theta_1$, $i = 2, \ldots, n$. Write $h_k(r_1, \ldots, r_n)$ as a polynomial in $r_1$:

$$
h_k(r_1, \ldots, r_n) = \sum_{i=0}^{k} h_i(r_2, \ldots, r_n) r_1^{k-i}.
$$

From (3.3), we have

$$
B_m(x) - \theta_1 = x - \theta_1 - (x - \theta_1) \frac{(x - \theta_1)^{m-2} h_{m-2}(r_1, \ldots, r_n)}{(x - \theta_1)^{m-1} h_{m-1}(r_1, \ldots, r_n)}
$$

$$
= (x - \theta_1) \left(1 - \frac{\sum_{i=0}^{m-2} h_i(r_2, \ldots, r_n)(x - \theta_1)^i}{\sum_{i=0}^{m-1} h_i(r_2, \ldots, r_n)(x - \theta_1)^i}\right)
$$

$$
= (x - \theta_1)^m \frac{h_{m-1}(r_2, \ldots, r_n)}{\sum_{i=0}^{m-1} h_i(r_2, \ldots, r_n)(x - \theta_1)^i}.
$$

Hence,

$$
\lim_{x \to \theta_1} \frac{B_m(x) - \theta_1}{(x - \theta_1)^m} = \lim_{x \to \theta_1} \frac{h_{m-1}(r_2, \ldots, r_n)}{1 + \sum_{i=1}^{m-1} h_i(r_2, \ldots, r_n)(x - \theta_1)^i}
$$

$$
= h_{m-1}(\theta_1 - \theta_2)^{-1}, \ldots, (\theta_1 - \theta_n)^{-1}).
$$

If $\theta_1$ is a root of multiplicity $s \geq 2$, without loss of generality we assume $\theta_1 = \theta_2 = \cdots = \theta_s$. Write $h_k(r_1, \ldots, r_n)$ as a polynomial in $r_1$ with coefficients that are polynomials in $r_{s+1}, \ldots, r_n$:

$$
h_k(r_1, \ldots, r_n) = \sum_{i=0}^{k} a_{k,i}(r_{s+1}, \ldots, r_n) r_1^{k-i}.
$$
The coefficient \( a_{k,i}(r_{x+1}, \ldots, r_n) \) is the product of \( h_i(r_{x+1}, \ldots, r_n) \) and the number of monomials in the form \( \theta_1^{k_1} \cdots \theta_s^{k_s} \), where \( k_1, \ldots, k_s \geq 0 \) and \( k_1 + \cdots + k_s = k - i \). Computing this number is a combinatorial problem, which is equivalent to counting the number of ways to place \( k - i \) identical balls and \( s - 1 \) identical bars in a row. Thus,

\[
a_{k,i}(r_{x+1}, \ldots, r_n) = \binom{k - i + s - 1}{s - 1} h_i(r_{x+1}, \ldots, r_n)
\]

for \( i = 0, \ldots, k \).

Again, from (3.3), we have

\[
B_m(x) - \theta_1 = x - \theta_1 - (x - \theta_1) \frac{(x - \theta_1)^{m-2} h_{m-2}(r_1, \ldots, r_n)}{(x - \theta_1)^{m-1} h_{m-1}(r_1, \ldots, r_n)}
\]

\[
= (x - \theta_1) \left( 1 - \frac{(m-x-3)}{s-1} + \sum_{i=1}^{m-2} a_{m-2,i}(r_{x+1}, \ldots, r_n)(x - \theta_1)^i \right).
\]

Hence,

\[
\lim_{x \to \theta_1} \frac{B_m(x) - \theta_1}{x - \theta_1} = 1 - \left( \frac{m+s-3}{s-1} \right) \left( \frac{m+s-2}{s-1} \right) = \frac{s-1}{s+m-2}.
\]

Interestingly, if all the roots of \( p \) are simple, there is another symmetric characterization of \( D_m(x) \) due to Schröder:

\[
D_m(x) = p(x)^{m+1} \sum_{i=1}^{n} \frac{x_i^{m+1}}{p'(\theta_i)}.
\]  

The bridge connecting identity (3.1) to identity (3.4) is a well-known formula in the theory of symmetric functions expressing a Schur function \( s_\lambda \) (which we will define shortly) as a polynomial in the complete symmetric functions \( h_r \):

\[
s_\lambda = \det(h_{\lambda_i-i+j})_{1 \leq i, j \leq n}.
\]

where \( \lambda \) is a partition of length \( n \), and \( h_r \) is defined to be zero for \( r < 0 \).

A Schur function in variables \( x_1, \ldots, x_n \), corresponding to the partition \( \lambda \) of length \( \leq n \) is defined as:

\[
s_\lambda(x_1, \ldots, x_n) = \frac{\det(x_1^{\lambda_j+n-j})_{1 \leq i, j \leq n}}{\det(x_1^{n-j})_{1 \leq i, j \leq n}}.
\]

Note on the right side of the above equation, both the numerator and denominator are skew-symmetric, so \( s_\lambda \) is symmetric. Furthermore, the denominator is a Vandermonde determinant and therefore is equal to \( \prod_{1 \leq i < j \leq n} (x_i - x_j) \). The numerator is divisible in
$\mathbb{Z}[x_1, \ldots, x_n]$ by each of the $(x_i - x_j)$ in the above product, hence $s_{\lambda} \in \Lambda_{n}^{[k]}$. A proof of identity (3.5) can be found in MacDonald [13, pp. 41–42].

Substituting $\lambda = (m)$ into (3.5), we get $s_{(m)} = h_m$. Thus,

$$h_m(r_1, \ldots, r_n) = s_{(m)}(r_1, \ldots, r_n) = \frac{\det(r_i^{\lambda_j+n-j})_{1 \leq i, j \leq n}}{\prod_{1 \leq i < j \leq n}(r_i - r_j)}.$$

Expanding $\det(r_i^{\lambda_j+n-j})_{1 \leq i, j \leq n}$ along the first column, and simplifying, we get

$$h_m(r_1, \ldots, r_n) = p(x) \sum_{i=1}^{n} \frac{r_i^{m+1}}{p'(\theta_i)},$$

which implies (3.4).

### 3.2. The multiple root variant of Basic Family

Corollary 3.4 shows that $B_m$ is an $m$th order method for approximating simple roots. For multiple roots, it has only linear convergence. To restore high order of convergence for multiple roots, we can apply $B_m$ to $p(x)/p'(x)$, whose roots are all simple and the same as those of $p(x)$. Surprisingly, this multiple root variant of the Basic Family has a recursive structure similar to that of the Basic Family.

**Proposition 3.5** (Jin and Kalantari [5]). For each integer $m \geq 0$, define

$$H_m(x) = \sum_{i=1}^{m} (-1)^{i-1} p(x)^{i-1} \frac{p^{(i)}(x)}{i!} H_{m-i}(x) + (-1)^m p(x)^m \frac{p^{(m+1)}(x)}{m!}. \tag{3.6}$$

Then, for each integer $m \geq 2$, the iteration function

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

is the application of $B_m$ to $p(x)/p'(x)$, hence has an order of convergence $m$ for all roots of $p(x)$.

Like the Basic Family, this multiple root variant also has a symmetric structure.

**Theorem 3.6.** For any integer $m \geq 0$,

$$H_m(x) = p(x)^{m+1} q_{m+1}(r_1, \ldots, r_n),$$

where $q_{m+1}$ is the $(m + 1)$th power sum.
Proof. The proof of this theorem is essentially the same as that of Theorem 3.2, except that here we use identity (2.7) instead of (2.5).

Now we have

$$F_m(x) = x - \frac{q_{m-1}(r_1, \ldots, r_n)}{q_m(r_1, \ldots, r_n)}.$$  \hspace{1cm} (3.7)

Corollary 3.7. Let \( s \) be the multiplicity of root \( \theta_1 \) of \( p(x) \). Without loss of generality, we assume \( \theta_1 = \theta_2 = \cdots = \theta_s \). Then,

$$\lim_{x \to \theta_1} \frac{F_m(x) - \theta_1}{(x - \theta_1)^m} = -\frac{1}{s} \sum_{i=s+1}^{n} (\theta_1 - \theta_i)^{-(m-1)}.$$  \hspace{1cm} (3.7)

Proof. The proof, which makes use of identity (3.7), is similar to that of Corollary 3.4.

This multiple root variant was also derived in Schröder [17], though through different means. But the iteration functions in the next subsection are new.

3.3. The Truncated Basic Family

According to the definition of the Basic Family, the \( m \)th order method \( B_m \) corresponding to polynomial \( p \) depends on \( p \) and its first \( m-1 \) derivatives. Traub [20] showed that the order of such a one-point root-finding iteration function is at most \( m \). So \( B_m \) has the highest possible order of convergence. If we restrict the use of derivatives of \( p \) up to the \( r \)th derivative when constructing \( B_{r+2} \) and up, then the order of methods in this “truncated” family is capped by \( t+1 \). In what follows, we show that these methods still achieve the optimal order of convergence, namely \( t+1 \), and furthermore, they share the same asymptotic error constant.

Definition 3.8 (The Truncated Basic Family of order \( t \)). For each integer \( t \geq 1 \), the Truncated Basic Family of order \( t \) is a family of iteration functions \( \{B_{m,t}(x)\}_{m=t+1}^{\infty} \), where \( B_{m,t}(x) \) is obtained from \( B_m(x) \) by replacing derivatives of \( p(x) \) of order higher than \( t \) by 0.

The Truncated Basic Family of order \( t \) also has a recursive definition: for each integer \( m \geq t+1 \), define

$$D_{m,t}(x) = \sum_{i=1}^{\min(m,t)} (-1)^{i-1} p(x)^{i-1} \frac{p^{(i)}(x)}{i!} D_{m-i,t}(x),$$  \hspace{1cm} (3.8)

where \( D_{0,t}(x) = 1 \). Then, for each integer \( m \geq t+1 \), the rational function \( B_{m,t}(x) \) is defined as

$$B_{m,t}(x) = x - p(x) \frac{D_{m-2,t}(x)}{D_{m-1,t}(x)}.$$  \hspace{1cm} (3.9)
For $t = 1$,

$$D_{m,1}(x) = \det \begin{bmatrix} p''(x) & 0 & 0 & \ldots & 0 \\ p(x) & p'(x) & 0 & \ddots & 0 \\ 0 & p(x) & p'(x) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \ldots & p'(x) \end{bmatrix}.$$ 

Thus, $D_{m,1}(x) = p'(x)^m$ and $B_{m,1}(x) = x - p(x)/p'(x) = B_2(x)$, that is, all members of the Truncated Basic Family of order 1 are identical to Newton’s method.

For $t = 2$,

$$D_{m,2}(x) = \det \begin{bmatrix} p'(x) & \frac{p''(x)}{2} & 0 & \ldots & 0 \\ p(x) & p'(x) & \frac{p''(x)}{2} & \ddots & 0 \\ 0 & p(x) & p'(x) & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \frac{p''(x)}{2} \\ 0 & 0 & 0 & \ldots & p'(x) \end{bmatrix}.$$ 

And $D_{m,2}(x)$ satisfies the following recurrence:

$$D_{0,2}(x) = 1, \quad D_{1,2}(x) = p'(x),$$

$$D_{m,2}(x) = p'(x)D_{m-1,2}(x) - \frac{1}{2} p(x)p''(x)D_{m-2,2}(x), \quad \forall m \geq 2.$$ 

The Truncated Basic Family of order 2 was studied and referred to as Halley’s Family in [7].

Analogous to the Basic Family and its multiple root variant, the Truncated Basic Family has a symmetric algebraic structure shown below.

**Theorem 3.9.** Let $t \geq 1$, $t \in \mathbb{Z}$, then, for each integer $s \geq 0$,

$$D_{s,t}(x) = p(x)^s g_{s,t}(r_1, \ldots, r_n), \quad (3.10)$$

where

$$g_{s,t}(r_1, \ldots, r_n) = h_s(r_1, \ldots, r_n) + \sum_{\sigma = n}^{\mid \lambda \mid = s} \sum_{t+1 \leq i(\lambda) \leq n} a_{s,t,\lambda} m_{i}(r_1, \ldots, r_n)$$
with \( a_{s,t,\lambda} \in \mathbb{Z} \), and \( h_s, m_s \) being complete and monomial symmetric functions, respectively. In particular, when \( s > t \), \( a_{s,t,(s'-(s-t))} = (-1)^{t+1}(s-t) \), where \((s'-(s-t))\) denotes the partition \(((s-t), 1, \ldots, 1)\).

**Proof.** The proof is by induction on \( s \).

When \( s \leq t \), \( D_{s,t}(x) = D_s(x) \), so from (3.1), we have

\[
D_{s,t}(x) = p(x)^t h_s(r_1, \ldots, r_n).
\]

Now assume (3.10) holds for \( s \leq k \), \( k \in \mathbb{Z}, k \geq t \), then for \( s = k+1 \), we have

\[
D_{k+1,t}(x) = \sum_{i=1}^{t} (-1)^{t-1} p(x)^{i-1} \frac{p^{(i)}(x)}{i!} D_{k+1-i,t}(x) = p(x)^{k+1} g_{k+1,t}(r_1, \ldots, r_n),
\]

where

\[
g_{k+1,t}(r_1, \ldots, r_n) = \sum_{i=1}^{t} (-1)^{t-1} e_i(r_1, \ldots, r_n) g_{k+1-i,t}(r_1, \ldots, r_n)
\]

\[
= \sum_{i=1}^{t} (-1)^{t-1} e_i(r_1, \ldots, r_n) h_{k+1-i}(r_1, \ldots, r_n)
\]

\[
+ \sum_{i=1}^{t} (-1)^{t-1} e_i(r_1, \ldots, r_n) \sum_{|\lambda|=k+1-i \atop t+1 \leq l(\lambda) \leq n} a_{k+1-i,t,\lambda} m_\lambda(r_1, \ldots, r_n)
\]

with \( a_{k+1-i,t,\lambda} \in \mathbb{Z} \) for \( i = 1, \ldots, t \).

From (2.5), we have

\[
\sum_{i=1}^{t} (-1)^{t-1} e_i h_{k+1-i} = h_{k+1} + \sum_{i=t+1}^{k+1} (-1)^{i} e_i h_{k+1-i}.
\]

Thus,

\[
g_{k+1,t}(r_1, \ldots, r_n) = h_{k+1}(r_1, \ldots, r_n)
\]

\[
= \sum_{i=t+1}^{k+1} (-1)^{i} e_i(r_1, \ldots, r_n) h_{k+1-i}(r_1, \ldots, r_n)
\]

\[
+ \sum_{i=1}^{t} (-1)^{t-1} e_i(r_1, \ldots, r_n) \sum_{|\lambda|=k+1-i \atop t+1 \leq l(\lambda) \leq n} a_{k+1-i,t,\lambda} m_\lambda(r_1, \ldots, r_n).
\]
It is easy to see that \( g_{k+1,t}(r_1, \ldots, r_n) - h_{k+1}(r_1, \ldots, r_n) \) is in \( A_n^{k+1} \), and consists of monomials in \( t+1 \) or more variables. Since \( m_\lambda \) such that \( l(\lambda) \leq n \) and \( |\lambda| = k+1 \) form a \( \mathbb{Z} \)-basis of \( A_n^{k+1} \), we have

\[
g_{k+1,t}(r_1, \ldots, r_n) - h_{k+1}(r_1, \ldots, r_n) = \sum_{|\lambda|=k+1 \atop t+1 \leq l(\lambda) \leq n} a_{k+1,t,\lambda} m_\lambda(r_1, \ldots, r_n),
\]

where \( a_{k+1,t,\lambda} \in \mathbb{Z} \).

If \( k = t \), (3.11) becomes

\[
g_{t+1,t}(r_1, \ldots, r_n) - h_{t+1}(r_1, \ldots, r_n) = (-1)^{t+1} e_{t+1}(r_1, \ldots, r_n),
\]

since the set \( \{ \lambda : |\lambda| = k+1 - i, k+1 \leq l(\lambda) \leq n, i \geq 1 \} \) is empty. By definition, \( e_{t+1} = m_{(t+1)} \). Thus, we have

\[
a_{t+1,t,(t+1)} = (-1)^{t+1},
\]

or equivalently,

\[
a_{k+1,t,(t+k+1-t)} = (-1)^{t+1}(k+1-t).
\]

If \( k > t \), the length of the partition \((t+1)(k+1-t)\) is \( t+1 \), so \( m_{(t+1)(k+1-t)}(r_1, \ldots, r_n) \) consists of monomials in \( t+1 \) variables, and in each of these monomials all variables except one have power 1. Multiplying out the right-hand side of (3.11), we get a sum of terms in \( A_n^{k+1} \). Expanding each term into a linear combination of basis variables \( m_\lambda \) such that \( l(\lambda) \leq n \) and \( |\lambda| = k+1 \), we find that only

\[
(-1)^{t+1} e_{t+1}(r_1, \ldots, r_n) h_{k-t}(r_1, \ldots, r_n)
\]

and

\[
a_{k,t,(t+k+1-t)} e_1(r_1, \ldots, r_n) m_{(t+k+1-t)}(r_1, \ldots, r_n)
\]

have non-zero \( m_{(t+k+1-t)}(r_1, \ldots, r_n) \) component. By induction hypothesis, we have

\[
a_{k,t,(t+k+1-t)} = (-1)^{t+1}(k-t).
\]

It is easy to see that

\[
e_{t+1}(r_1, \ldots, r_n) h_{k-t}(r_1, \ldots, r_n) = m_{(t+k+1-t)}(r_1, \ldots, r_n) + L_1(r_1, \ldots, r_n),
\]

\[
e_1(r_1, \ldots, r_n) m_{(t+k+1-t)}(r_1, \ldots, r_n) = m_{(t+k+1-t)}(r_1, \ldots, r_n) + L_2(r_1, \ldots, r_n),
\]

where \( L_1 \) and \( L_2 \) are integral linear combinations of \( m_\lambda \) such that \( |\lambda| = k+1, t+1 \leq l(\lambda) \leq n, \lambda \neq (t+k+1-t) \).

Thus,

\[
a_{k+1,t,(t+k+1-t)} = (-1)^{t+1} + (-1)^{t+1}(k-t) = (-1)^{t+1}(k+1-t).
\]

So (3.10) holds for \( s = k+1 \). And by induction, (3.10) holds for all \( s \geq 0 \). \( \square \)
Using Theorem 3.9, it is easy to derive the convergence property for the Truncated Basic Family.

We first observe that for \( t = 1 \), all members of the Truncated Basic Family of order 1 are identical to Newton’s method. And for each \( t \geq 2 \), we have \( B_{t+1,t} = B_t \), which is a \((t + 1)\)th order method. The following corollary settles the orders of all other Truncated Basic Family methods.

**Corollary 3.10.** Let \( t \geq 2 \), \( t \in \mathbb{Z} \), then for each integer \( s \geq t + 2 \), \( B_{s,t}(x) \) has an order of convergence \( t + 1 \) for a simple root of \( p \), say \( \theta_1 \), and

\[
\lim_{x \to \infty} \frac{B_{s,t}(x) - \theta_1}{(x - \theta_1)^{t+1}} = \frac{(-1)^{t+1} p^{(t+1)}(\theta_1)}{(t + 1)! p'(\theta_1)}.
\]

**Proof.** From (3.9) and (3.10), we have

\[
B_{s,t}(x) = x - \frac{g_{s-2,t}(r_1, \ldots, r_n)}{g_{s-1,t}(r_1, \ldots, r_n)},
\]

where

\[
g_{u,t}(r_1, \ldots, r_n) = h_u(r_1, \ldots, r_n) + \sum_{\|\lambda\| = u \atop t+1 \leq l(\lambda) \leq n} a_{u,t,\lambda} m_{\lambda}(r_1, \ldots, r_n).
\]

Now write \( g_{u,t}(r_1, \ldots, r_n), u \geq t \) as a polynomial in \( r_1 \):

\[
g_{u,t}(r_1, \ldots, r_n) = \sum_{i=0}^{n} h_i(r_2, \ldots, r_n) r_1^{u-i} + (-1)^{t+1}(u-t)e_i(r_2, \ldots, r_n)r_1^{u-t}
\]

\[
+ \sum_{i=t+1}^{n} c_i(r_2, \ldots, r_n)r_1^{u-i},
\]

where \( c_i \) are polynomials in \( n - 1 \) variables. We have used the facts that, when \( u > t \), the biggest part of \( \lambda \) such that \( |\lambda| = u \) and \( t + 1 \leq l(\lambda) \leq u \) is at most \( u - t \), which only appears in \((1^t(u-t))\), and \( a_{u,t,1^t(u-t)} = (-1)^{t+1}(u-t) \) when \( u > t \).

Thus, for \( s \geq t + 2 \), we have

\[
B_{s,t}(x) - \theta_1 = x - \theta_1 - (x - \theta_1) \frac{(x - \theta_1)^{t-2} g_{s-2,t}(r_1, \ldots, r_n)}{(x - \theta_1)^{t-1} g_{s-1,t}(r_1, \ldots, r_n)}
\]

\[
= (x - \theta_1) \left( 1 - \frac{(x - \theta_1)^{t-2} g_{s-2,t}(r_1, \ldots, r_n)}{(x - \theta_1)^{t-1} g_{s-1,t}(r_1, \ldots, r_n)} \right)
\]

\[
= (x - \theta_1)^{t+1} \frac{(-1)^{t+1} e_t(r_2, \ldots, r_n) + Q_1(r_1^{-1}, r_2, \ldots, r_n)(x - \theta_1)}{1 + Q_2(r_1^{-1}, r_2, \ldots, r_n)(x - \theta_1)},
\]

where \( Q_1 \) and \( Q_2 \) are polynomials in \( n - 1 \) variables.

The proof is complete.
where $Q_1$ and $Q_2$ are polynomials in $n$ variables.

Hence,

$$\lim_{x \to \theta_1} B_{s,t}(x) - \theta_1 = (-1)^{t+1} e_t((\theta_1 - \theta_2)^{-1}, \ldots, (\theta_1 - \theta_n)^{-1})$$

$$= \frac{(-1)^{t+1} \rho(t+1)(\theta_1)}{(t+1)! \rho'(\theta_1)}, \quad \square$$

4. A recipe for constructing new high order methods

All of the root-finding iteration functions we have studied in this paper possess the following symmetric structure:

$$I_t(x) = x - \frac{SF_{t-2}(r_1, \ldots, r_n)}{SF_{t-1}(r_1, \ldots, r_n)}, \quad (4.1)$$

where $t \in \mathbb{Z}$, $t \geq 2$, and $\{SF_i\}_{i=0}^\infty$ is a sequence of symmetric functions with integer coefficients. Recall that $r_j = 1/(x - \theta_j)$, for $j = 1, \ldots, n$, and $\theta_1, \ldots, \theta_n$ are complex roots of $p$.

Conversely, any iteration function $I_t(x)$ defined by (4.1) is a rational function in $x$, $p(x)$ and its derivatives. To see this, we first state the fundamental theorem of symmetric functions.

**Theorem 4.1** (The fundamental theorem of symmetric functions). Any symmetric function with integer coefficients is uniquely expressible as a polynomial in the elementary symmetric functions $e_1, e_2, e_3, \ldots$ with integer coefficients.

The proof of this theorem can be found in [13]. The fundamental theorem of symmetric functions says that $SF_t$ can always be represented by a polynomial in $e_1, e_2, e_3, \ldots$, which, in turn, reduce to rational functions in $p(x)$ and its derivatives by Lemma 3.1.

The Basic Family and its multiple root variant are obtained by letting $SF_i$ be the $i$th complete symmetric function $h_i$ and $(i+1)$th power sum $q_{i+1}$, respectively. New high order root-finding algorithms can be “invented” by picking an appropriate sequence of symmetric functions $\{SF_i\}_{i=0}^\infty$. Among many promising candidates for $SF_i$ are simple monomial symmetric functions. The eligibility of a candidate sequence can be checked using the technique employed in the proof of Corollary 3.4. For example, we may try $SF_i = m_{(i+2,1)}$, where $m_{(i+2,1)}$ denotes the monomial symmetric function corresponding to the partition $(i+2, 1)$. It is easy to verify that

$$I_t(x) = x \frac{m_{(t,1)}(r_1, \ldots, r_n)}{m_{(t+1,1)}(r_1, \ldots, r_n)}$$

has local $t$th order of convergence to a simple root of $p$ and local $(t+1)$th order of convergence to a multiple root of $p$, which is a bit unusual.
We now derive the first member of this family:

\[ I_2(x) = x - \frac{m_{(2,1)}(r_1, \ldots, r_n)}{m_{(3,1)}(r_1, \ldots, r_n)}. \]

First, we use Stembridge’s Maple package for symmetric functions, SF [19], to convert \( m_{(2,1)} \) and \( m_{(3,1)} \) into polynomials in the elementary symmetric functions \( e_1, e_2, e_3, \ldots \):

\[
m_{(2,1)} = e_1 e_2 - 3e_3, \quad m_{(3,1)} = e_1^2 e_2 - e_1 e_3 - 2e_2^2 + 4e_4.
\]

Then, replacing \( e_i \) by \( p^{(i)}(x)/(i! p(x)) \) (Lemma 3.1) and simplifying, we get

\[
I_2(x) = x - \frac{p(x) - 3p(x) p'(x) - 3p(x) p''(x) + p(x)^2 p'''(x) + p(x) p'(x)p''(x) - 3p(x) p''(x) - 2p'(x)^3}{3p'(x) + 2p''(x)}.
\]

which is a 2nd order method for simple roots and 3rd order method for multiple roots.

Like the Basic Family, this family also has recursive and determinantal formulae in terms of \( p(x) \) and its derivatives. We will develop them in a separate note.

Computationally, new families of root-finding algorithms such as the above one are usually less efficient than the Basic Family, therefore are not as practical. Nevertheless they showcase the rich variety of high order methods and demonstrate the close connection between symmetric function theory and root-finding algorithms. Visualization of the dynamics of these iteration functions, called Polynomiography in [12], could offer new insights into theoretical frontiers, and provide a new medium for art and education as well.

Many individual rational root-finding iteration functions have the symmetric structure characterized by (4.1) as well. A prominent example is McMullen’s generally convergent algorithm for cubics (see [14]):

\[
T(x) = x - \frac{3p(x) p''(x) - 2p'(x)^3}{4p(x) p'(x) p''(x) - 3p(x)^2 p'''(x) - 2p'(x)^3}.
\]

It would be interesting to investigate whether this 3rd order algorithm can be extended to an infinite family of high order generally convergent algorithms for cubics, by picking an appropriate sequence of \( \{SF_i\}_{i=0}^{\infty} \) for (4.1).

Acknowledgment

The first author thanks Professor Doron Zeilberger for kindly introducing him to the beautiful theory of symmetric functions and the art of Maple programming.

References