Journal of Computational and Applied Mathematics 321 (2017) 261–269

Contents lists available at ScienceDirect

Journal of Computational and Applied Mathematics

journal homepage: www.elsevier.com/locate/cam

General approach to constructing optimal multipoint families of iterative methods using Hermite's rational interpolation☆

Jovana Džunić*, Ivan Damnjanović

Faculty of Electronic Engineering, University of Niš, 18 000 Niš, Serbia

ARTICLE INFO

Article history:
Received 11 October 2016
Received in revised form 23 February 2017

MSC:
65H05

Keywords:
Iterative methods
Nonlinear equations
Hermite rational interpolation
Divided difference

ABSTRACT

We discuss accelerating convergence of multipoint iterative methods for solving scalar equations, using particular type rational interpolant. Both derivative-free and Newton-type methods are investigated simultaneously. As a conclusion a Theorem of König's type for multipoint iterations is stated. A new optimal multipoint family of methods based on rational interpolation is constructed. The iteration uses n function evaluations per cycle and $\mathcal{O}(j)$ operations in jth step of a single iteration to obtain 2^{n-1} order of convergence. Several equivalent forms of the obtained iterates and development techniques are presented.

1. Introduction

Multipoint methods are among the most efficient tools in approximating a single root of a nonlinear equation by iterative means. The search for reliable methods to construct these powerful root-finders is very intensive and has increased over the years. For details and relevant references, the reader is referred to [1–3] or [4]. There is undoubtedly a tight relation between function interpolation and iterative root-finders. The aim of this communication is to investigate the use of a special type of rational interpolation in the construction of multipoint methods.

The first optimal two-point method was constructed by Ostrowski [5] in 1960. For the approximate solution of a nonlinear equation $f(x) = 0$, Ostrowski derived the two-point method

$$\begin{cases} y_k = x_k - \dfrac{f(x_k)}{f'(x_k)}, \\ x_{k+1} = x_k - \dfrac{f(y_k) - f(x_k)}{2f(y_k) - f(x_k)} \cdot \dfrac{f(x_k)}{f'(x_k)}, \\ \qquad\quad = y_k - \dfrac{f(x_k)}{f(x_k) - 2f(y_k)} \cdot \dfrac{f(y_k)}{f'(x_k)}, \end{cases} \tag{1}$$

using interpolation of the function f through points x_k and y_k by a rational function $w(t) = \frac{t-a}{bt-c}$, with a linear numerator. This numerator structure enables direct evaluation of a unique root of w. The root is taken as the next approximation of the two-point iterative scheme (1). Newton's method is used as a pre-conditioner. We follow Ostrowski's idea to define new

☆ This research was supported by the Serbian Ministry of Education and Science under grant number 174022.
* Corresponding author.
E-mail address: jovana.dzunic@gmail.com (J. Džunić).

http://dx.doi.org/10.1016/j.cam.2017.02.039
0377-0427/

optimal n-point families of methods, with and without derivatives. This approach can also be used to double the rate of convergence of any optimal multipoint scheme at the cost of one new function evaluation per iteration. Methods of order eight and sixteen with Newton's pre-conditioner were recently designed in this manner in [6,7]. One-point methods with memory of this type were discussed in [8–10], etc. This communication aims to give a unified approach for both derivative-free and Newton-type multipoint methods based on the specific type rational interpolant. Several equivalent forms and ways to derive formulas are presented.

In this communication we are mainly concerned with the theoretical aspects when constructing optimal root-finding algorithms. A general method which provides reliable means of convergence upgrade is presented. Particular methods of such type, with practical relevance, have been developed in earlier research and we cite them here and give explicit formulas at the end. For details on numerical test results the reader is referred to particular publications.

Relying on polynomial interpolation in the development of iterative methods has been long recognized as a fruitful technique. Spurred by earlier research examples, in this paper we explore a different type of interpolation, namely with rational functions. It is sometimes encountered that modelling by rational functions is superior to polynomial approximation. The key issue is the ability to model functions with poles. Quite often polynomials exhibit poor interpolating properties due to the presence of poles in an analytic continuation of the function being modelled. Such poles can ruin a polynomial approximation, even when restricted to real valued arguments. A rational function approximation, on the other hand, will show good approximating properties as long as it has a high enough polynomial degree in its denominator. For details of the mentioned characteristics of the described approximations see, e.g., [11,12]. For these reasons we develop a strategy for multipoint iterations based on rational interpolation.

The paper is organized as follows. In the succeeding section we give basic definitions and properties regarding the topic to be explored. In Section 3 we use a rational interpolant to increase the order of convergence of an arbitrary optimal scheme, resulting in an efficiency advancement. In Section 4 convergence properties are proven. In the final section the same approach is employed to design a new optimal n-point family of iterative methods.

2. Preliminaries

Let a class $\mathcal{F}$ be defined by

$$\mathcal{F} = \{f \,|\, f \text{ is an analytic function defined on an open connected set } D_f \subset \mathbb{C} \text{ which contains a simple zero } \alpha_f \text{ of } f \text{ and } f' \text{ does not vanish on } D_f\}.$$

Definition 2.1. Let φ be an iteration function and α_f denote a zero of the function f. If there exists an $r = r(\varphi) \in \mathbb{N}$ such that for any $f \in \mathcal{F}$,

$$\lim_{x\to\alpha_f} \frac{|\varphi(f)(x) - \alpha_f|}{|x-\alpha_f|^r} = A \tag{2}$$

exists for a constant A which does not vanish at least for one $f \in \mathcal{F}$, then φ is said to have the *order of convergence* r. The constant $A > 0$ is called an *asymptotic error constant* of iteration φ.

Definition 2.1 of order of convergence will suffice for multipoint methods without memory, which are the topic of this communication.

We consider the case where branch cuts of an analytic function are not taken in a neighbourhood of the sought zero α. It is known that root-finding algorithms are sensitive to the presence of branch points. This issue is not the topic of our communication, still we point to a recent elegant idea that was proposed in [13] to overcome this problem.

The goal in constructing iterative methods for solving nonlinear equations is to attain as high as possible order of convergence with the minimal computational cost. It is closely related to terms of efficiency and optimality of iteration functions. Introduce the notion of *information volume* as the number of new function evaluations (counting derivatives as well). Comparison between iterative methods can be carried out using efficiency index. The *efficiency index* of the iteration φ with order of convergence r and volume of information n is defined with $EI(\varphi) = r^{1/n}$. This is a very reduced efficiency measure since it neglects the computational cost of the very iterative formula. Many authors in recent publications are deceived by this simplification and the only aim appointed is to achieve as high as possible order of convergence. This is particularly characteristic and hazardous for methods with memory that use a large amount of self-correcting parameters. In this paper we make count of a number of rational operations per step in addition to the information volume.

First formal statement of limitation regarding efficiency of multipoint methods dates back to 1974. In their fundamental paper [14], H.T. Kung and J.F. Traub stated the following hypothesis.

Kung–Traub's conjecture: Multipoint iterative methods without memory, costing $n+1$ function evaluations per iteration, have order of convergence at most 2^n.

Iterations without memory using $n+1$ function evaluations and order 2^n are called optimal. Optimal methods have the efficiency index $EI(\varphi) = 2^{(n-1)/n}$.

When designing multipoint methods, we are dealing with clustered sets of points. Loss of significant figures and instability of calculation often occur in such scenarios. Multi-precision arithmetic and numerically stable formulas help overcome these issues. Divided differences scheme shows good numerical behaviour with clustered data. For this reason we use them in the development techniques for multipoint iterations. In the sequel we recall definition and some basic properties of divided differences [15,16].

For $x_i \neq x_j$ when $i \neq j$ divided differences are defined as

$$f[x_0] = f(x_0),$$
$$f[x_0, x_1] = \frac{f[x_0] - f[x_1]}{x_0 - x_1},$$
$$f[x_0, x_1, \ldots, x_n] = \frac{f[x_0, \ldots, x_{n-1}] - f[x_1, \ldots, x_n]}{x_0 - x_n}.$$

Properties of divided differences that are of relevance:

Linearity:

$$(f + h)[x_0, x_1, \ldots, x_n] = f[x_0, x_1, \ldots, x_n] + h[x_0, x_1, \ldots, x_n],$$
$$(\lambda \cdot f)[x_0, x_1, \ldots, x_n] = \lambda \cdot f[x_0, x_1, \ldots, x_n].$$

Symmetry:

$$f[x_0, x_1, \ldots, x_n] = f[x_{i_0}, x_{i_1}, \ldots, x_{i_n}],$$

where $i_0, i_1, \ldots, i_n$ is any permutation of the set $\{0, 1, \ldots, n\}$.

Leibnitz rule:

$$(f \cdot h)[x_0, x_1, \ldots, x_n] = \sum_{j=0}^{n} f[x_0, \ldots, x_j] h[x_j, \ldots, x_n].$$

The last formula is in fact due to Popoviciu [17,18] and Steffensen [19], but since it generalizes the nth derivative of the product of functions it is usually credited to Leibnitz.

Due to continuity, the definition of divided differences can be generalized to cases when some or all x_j coincide.

$$f[x_0, x_0] = f'(x_0), \qquad f[\underbrace{x_0, \ldots, x_0}_{m+1}] = \frac{f^{(m)}(x_0)}{m!}.$$

The above listed properties remain valid with this definition of divided differences in a broader sense.

3. Accelerating convergence

We restrict our study to solving a scalar nonlinear equation of the form $f(x) = 0$ by iterative means, where the sought zero α is simple. The multipoint iteration function of the form $x_{k+1} = \varphi(x_k),\ k \in \mathbb{N}$, is investigated for the close enough initial value x_0. Auxiliary values within a single multipoint iteration will be denoted with $y_j = y_j(k),\ j = 0, \ldots, n$, with the iteration counter k omitted. It is assumed that $y_0 = x_k,\ y_1 = y_0 + \gamma f(y_0),\ \gamma \in \mathbb{C}$, and $x_{k+1} = y_n$. Note that $\gamma = 0$ is allowed. In this manner both Steffensen-type methods and Newton-type iterations are explored simultaneously.

The following multipoint optimal scheme is considered

$$\begin{cases} y_0 = x_k, \qquad y_1 = y_0 + \gamma f(y_0), \\ y_2 = y_0 - \dfrac{f(y_0)}{f[y_1, y_0]}, \\ y_j = \Psi_j(y_0, y_1, \ldots, y_{j-1}), \quad j = 3, 4, \ldots, n-1, \\ x_{k+1} = y_{n-1}, \end{cases} \tag{3}$$

where each $y_j = \Psi_j(y_0, y_1, \ldots, y_{j-1}),\ j \geq 2$, is an optimal j-point scheme of order 2^{j-1} using j evaluations of f. Obviously, the approximation y_2 in (3) defines Newton's step for $\gamma = 0$ and Traub-Steffensen's iteration [1,20] when $\gamma \neq 0$.

The following two sampling cases are simultaneously considered:

1° for $\gamma = 0$ information used: $f(y_0), f'(y_0), f(y_2), \ldots, f(y_{n-2})$;
2° for $\gamma \neq 0$ information used: $f(y_0), f(y_1), f(y_2), \ldots, f(y_{n-2})$.

Both sampling cases have been proved to provide best efficiency for multipoint methods, see [14]. Optimal iteration scheme of the form (3) with sampling cases 1^o or 2^o will be denoted *NS*-iteration, after Newton and Steffensen.

Introduce Hermite polynomial basis: $B_{0,n}(t) = 1$,

$$B_{j,n}(t) = (t - y_{n-j}) B_{j-1,n}(t), \quad j = 1, \ldots, n.$$

We thus write [15,16]

$$F(t) = \sum_{j=0}^{n-1} F[y_{n-1}, \dots, y_{n-1-j}]B_{j,n}(t) + F[y_{n-1}, \dots, y_0, t]B_{n,n}(t), \tag{4}$$

for any function F, where for the summation index value $j = 0$ the addend equals $F[y_{n-1}]$.

The order of convergence of the scheme (3) will be increased by an additional step using one more function evaluation $f(y_{n-1})$. A new approximation y_n of the sought zero α will be calculated as the unique zero of a rational function

$$w_n(t) = \frac{P_n(t)}{Q_n(t)} = \frac{t - y_n}{\sum_{j=0}^{n-2} a_j B_{j,n}(t)}, \tag{5}$$

where the coefficients y_n and a_j take such values that the interpolation conditions

$$w_n(y_j) = f(y_j),\ j = 1, \dots, n-1, \quad \text{and} \quad w_n[y_1, y_0] = f[y_1, y_0], \tag{6}$$

are satisfied.

Conditions (6) can be transformed into a set of equations

$$\left(Q_n - \frac{P_n}{f}\right)[y_{n-1}, \dots, y_j] = 0, \quad j = 0, \dots, n-1, \tag{7}$$

where for $j = n-1$ the above condition takes form $\left(Q_n - \frac{P_n}{f}\right)[y_{n-1}] = 0$. Since

$$\begin{aligned}
&P_n[y_{n-1}, y_{n-2}] = 1,\\
&P_n[y_{n-1}, \dots, y_j] = 0, \quad \text{for } 0 \le j < n-2,\\
&Q_n[y_{n-1}, \dots, y_{n-1-j}] = a_j, \quad \text{for } 0 \le j < n-1,\\
&Q_n[y_{n-1}, \dots, y_0] = 0,
\end{aligned}$$

and by the Leibniz rule for divided differences, conditions (7) in matrix form read

$$\begin{bmatrix}
\frac{1}{f}[y_{n-1}] & 1 & 0 & \cdots & 0\\
\frac{1}{f}[y_{n-1}, y_{n-2}] & 0 & 1 & \cdots & 0\\
\frac{1}{f}[y_{n-1}, \dots, y_{n-3}] & 0 & 0 & \cdots & 0\\
 & & \cdots & & \\
\frac{1}{f}[y_{n-1}, \dots, y_1] & 0 & 0 & \cdots & 1\\
\frac{1}{f}[y_{n-1}, \dots, y_0] & 0 & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix} y_n\\ a_0\\ a_1\\ \vdots\\ a_{n-3}\\ a_{n-2} \end{bmatrix}
=
\begin{bmatrix}
y_{n-1}\frac{1}{f}[y_{n-1}]\\
y_{n-1}\frac{1}{f}[y_{n-1}, y_{n-2}] + \frac{1}{f}[y_{n-2}]\\
y_{n-1}\frac{1}{f}[y_{n-1}, \dots, y_{n-3}] + \frac{1}{f}[y_{n-2}, y_{n-3}]\\
\vdots\\
y_{n-1}\frac{1}{f}[y_{n-1}, \dots, y_1] + \frac{1}{f}[y_{n-2}, \dots, y_1]\\
y_{n-1}\frac{1}{f}[y_{n-1}, \dots, y_0] + \frac{1}{f}[y_{n-2}, \dots, y_0]
\end{bmatrix}.$$

The last row of the matrix reduces to

$$y_n = y_{n-1} + \frac{\frac{1}{f}[y_{n-2}, \dots, y_0]}{\frac{1}{f}[y_{n-1}, \dots, y_0]}. \tag{8}$$

Remark 3.1. Note that Newton's and Traub-Steffensen's methods are both of the form (8).

From the table of divided differences and formula (8) it is easily estimated that the new approximation requires $\mathcal{O}(n^2)$ rational operations in addition to one new function evaluation.

We explore some equivalent forms of the formula (8). Namely, with a change of polynomial basis in the denominator Q_n of the rational function w_n, we can obtain different appearance of the investigated iteration.

By reordering points in the Hermite polynomial basis we obtain the first equivalent formula

$$y_n = y_0 + \frac{\frac{1}{f}[y_1, \dots, y_{n-1}]}{\frac{1}{f}[y_0, \dots, y_{n-1}]}.$$

If we use Taylor's polynomial basis $T_j(t) = (t - y_0)^j,\ j \in \mathbb{N}$, we can rewrite

$$Q_n(t) = \sum_{j=0}^{n-2} q_j T_j(t),$$

for some coefficients $q_j \in \mathbb{C}$. It is convenient to consider interpolating conditions (6) in the form

$$(fQ_n - P_n)[y_0] = 0, \qquad (fQ_n - P_n)[y_k, y_0] = 0, \quad k = 1, 2 \ldots, n-1,$$

or equivalently

$$\begin{bmatrix} 1 & f(y_0) & 0 & 0 & \cdots & 0 \\ 0 & f[y_1, y_0] & f(y_1) & f(y_1)(y_1 - y_0) & \cdots & f(y_1)(y_1 - y_0)^{n-3} \\ 0 & f[y_2, y_0] & f(y_2) & f(y_2)(y_2 - y_0) & \cdots & f(y_2)(y_2 - y_0)^{n-3} \\ & & & \cdots & & \\ 0 & f[y_{n-1}, y_0] & f(y_{n-1}) & f(y_{n-1})(y_{n-1} - y_0) & \cdots & f(y_{n-1})(y_{n-1} - y_0)^{n-3} \end{bmatrix} \begin{bmatrix} y_n \\ q_0 \\ q_1 \\ \vdots \\ q_{n-2} \end{bmatrix} = \begin{bmatrix} y_0 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}.$$

Determinants for the Cramer's rule read:

$$\begin{aligned} \Delta &= \begin{vmatrix} f[y_1, y_0] & f(y_1) & f(y_1)(y_1 - y_0) & \cdots & f(y_1)(y_1 - y_0)^{n-3} \\ f[y_2, y_0] & f(y_2) & f(y_2)(y_2 - y_0) & \cdots & f(y_2)(y_2 - y_0)^{n-3} \\ & & \cdots & & \\ f[y_{n-1}, y_0] & f(y_{n-1}) & f(y_{n-1})(y_{n-1} - y_0) & \cdots & f(y_{n-1})(y_{n-1} - y_0)^{n-3} \end{vmatrix} \\ &= \sum_{j=1}^{n-1} (-1)^{j+1} f[y_j, y_0] \Big(\prod_{\substack{i=1 \\ i \neq j}}^{n-1} f(y_i) \Big) V(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_{n-1}) \\ &= \sum_{j=1}^{n-1} f[y_j, y_0] \Big(\prod_{\substack{i=1 \\ i \neq j}}^{n-1} \frac{f(y_i)}{y_i - y_j} \Big) V(y_1, \ldots, y_n), \end{aligned}$$

where the symbol V stands for the Vandermonde determinant.

$$\begin{aligned} \Delta(y_n) &= \begin{vmatrix} y_0 & f(y_0) & 0 & 0 & \cdots & 0 \\ 1 & f[y_1, y_0] & f(y_1) & f(y_1)(y_1 - y_0) & \cdots & f(y_1)(y_1 - y_0)^{n-3} \\ 1 & f[y_2, y_0] & f(y_2) & f(y_2)(y_2 - y_0) & \cdots & f(y_2)(y_2 - y_0)^{n-3} \\ & & & \cdots & & \\ 1 & f[y_{n-1}, y_0] & f(y_{n-1}) & f(y_{n-1})(y_{n-1} - y_0) & \cdots & f(y_{n-1})(y_{n-1} - y_0)^{n-3} \end{vmatrix} \\ &= y_0 \Delta - f(y_0) \sum_{j=1}^{n-1} (-1)^{j+1} \Big(\prod_{\substack{i=1 \\ i \neq j}}^{n-1} f(y_i) \Big) V(y_1, \ldots, y_{j-1}, y_{j+1}, \ldots, y_{n-1}) \\ &= y_0 \Delta - f(y_0) \sum_{j=1}^{n-1} \Big(\prod_{\substack{i=1 \\ i \neq j}}^{n-1} \frac{f(y_i)}{y_i - y_j} \Big) V(y_1, \ldots, y_n). \end{aligned}$$

We conclude

$$y_n = \frac{\Delta(y_n)}{\Delta} = y_0 - \frac{\sum_{j=1}^{n-1} \prod_{\substack{i=1 \\ i \neq j}}^{n-1} \frac{f(y_i)}{y_i - y_j}}{\sum_{j=1}^{n-1} f[y_j, y_0] \prod_{\substack{i=1 \\ i \neq j}}^{n-1} \frac{f(y_i)}{y_i - y_j}} f(y_0). \tag{9}$$

Formula (9) can be regarded as the 'barycentric' form of iterative formula (8). Such form is known to be computationally robust and very reliable [21,22].

Another equivalent form of (8) can be obtained using expanded form of divided differences of the reciprocal function [23]

$$\frac{1}{f}[y_0, \ldots, y_s] = \sum_{j=1}^{s} (-1)^j \sum_{0 = i_0 < \cdots < i_j = s} \frac{\prod_{l=0}^{j-1} f[y_{i_l}, \ldots, y_{i_{l+1}}]}{f(y_{i_0}) f(y_{i_1}) \cdots f(y_{i_j})}.$$

4. Convergence theorem

In this section we show that the presented approach of rational function interpolation provides the desired optimal convergence order 2^{n-1}.

Theorem 4.1. *Let $f(x) \in \mathcal{F}$ and define the iteration function*

$$\begin{cases} y_0 = x_k, \qquad y_1 = y_0 + \gamma f(y_0), \\ y_j = \Psi_j(y_0, y_1, \dots, y_{j-1}), \quad j = 2, 3, \dots, n-1, \\ x_{k+1} = y_n = y_{n-1} + \dfrac{\frac{1}{f}[y_{n-2}, \dots, y_0]}{\frac{1}{f}[y_{n-1}, \dots, y_0]}, \end{cases} \tag{10}$$

where each $y_j = \Psi_j(y_0, y_1, \dots, y_{j-1})$, $2 \le j \le n-1$, is an optimal NS-iteration of order 2^{j-1}. Then the scheme (10) is an optimal NS-iteration and attains 2^{n-1} order of convergence.

Proof. Errors of approximations to the sought zero α obtained in a single iteration cycle will be labelled $\varepsilon_j = \alpha - y_j$, $j = 0, \dots, n$, with the iteration counter k omitted. According to the construction of (10) we have

$$\varepsilon_j = \mathcal{O}\big(\varepsilon_0^{2^{j-1}}\big), \quad j = 1, \dots, n-1. \tag{11}$$

Introduce functions $\rho(t) = \frac{t-\alpha}{f(t)}$ and $R(t) = P_n(t)\rho(t) - (t-\alpha)Q_n(t)$, where P_n and Q_n satisfy (7). Since α is an isolated zero of an analytic function $f \in \mathcal{F}$ it follows that $\rho(t)$ is analytic within some Jordan curve $\gamma \subset D_f$, enclosing the points y_j and bounded on γ. Also, $\rho(\alpha) \neq 0$.

For all $j = 0, \dots, n-1$ we have

$$R[y_{n-1}, \dots, y_j] = \left(P_n\rho - (\cdot - \alpha)\frac{P_n}{f}\right)[y_{n-1}, \dots, y_j] = 0.$$

Thus, $(t-\alpha)Q_n(t)$ is the unique interpolating polynomial of minimal degree of the function $\varrho(t) = P_n(t)\rho(t)$ and $R(t)$ is the error function of the interpolation. The well known remainder relation of Hermite's interpolating polynomial (4) for $t = \alpha$ reads

$$R(\alpha) = \varrho[y_{n-1}, \dots, y_0, \alpha] \cdot B_{n,n}(\alpha). \tag{12}$$

Since $\varrho(t) = (t-\alpha)\rho(t) + (\alpha - y_n)\rho(t)$, using divided differences properties it follows

$$\varrho[y_{n-1}, \dots, y_0, \alpha] = \rho[y_{n-1}, \dots, y_0] + (\alpha - y_n)\rho[y_{n-1}, \dots, y_0, \alpha]. \tag{13}$$

Substituting (13) into (12) yields

$$R(\alpha) = \big(\rho[y_{n-1}, \dots, y_0] + P_n(\alpha)\rho[y_{n-1}, \dots, y_0, \alpha]\big) \cdot B_{n,n}(\alpha).$$

From the definition of $R(t)$ we have

$$R(\alpha) = P_n(\alpha)\rho(\alpha).$$

Therefore, having in mind $P_n(\alpha) = \varepsilon_n$, it follows

$$\varepsilon_n = \frac{\rho[y_{n-1}, \dots, y_0] \cdot \prod_{j=0}^{n-1} \varepsilon_j}{\rho(\alpha) - \rho[y_{n-1}, \dots, y_0, \alpha] \cdot \prod_{j=0}^{n-1} \varepsilon_j} = \mathcal{O}(\varepsilon_0^{2^{n-1}}). \tag{14}$$

To obtain an asymptotic error constant of iteration (10), we rewrite assumptions (11) in the form

$$\lim_{\varepsilon_0 \to 0} \frac{|y_j - \alpha|}{|y_0 - \alpha|^{2^{j-1}}} = \lim_{\varepsilon_0 \to 0} \frac{|\varepsilon_j|}{|\varepsilon_0|^{2^{j-1}}} = A_j. \tag{15}$$

From (14) we find

$$\lim_{\varepsilon_0 \to 0} \frac{\varepsilon_n}{\prod_{j=0}^{n-1} \varepsilon_j} = \lim_{\varepsilon_0 \to 0} \frac{\rho[y_{n-1}, \dots, y_0]}{\rho(\alpha) - \rho[y_{n-1}, \dots, y_0, \alpha] \cdot \prod_{j=0}^{n-1} \varepsilon_j}$$

$$= \frac{\rho^{(n-1)}(\alpha)}{(n-1)!\rho(\alpha)} = \frac{f'(\alpha)\rho^{(n-1)}(\alpha)}{(n-1)!}. \tag{16}$$

It follows from (15) and (16)

$$\lim_{\varepsilon_0\to 0}\frac{|\varepsilon_n|}{|\varepsilon_0|^{2^{n-1}}} = \lim_{\varepsilon_0\to 0}\frac{|\varepsilon_n|}{\prod_{j=0}^{n-1}|\varepsilon_j|}\cdot\frac{\prod_{j=0}^{n-1}|\varepsilon_j|}{|\varepsilon_0|^{2^{n-1}}} = \frac{|f'(\alpha)\rho^{(n-1)}(\alpha)|}{(n-1)!}\prod_{j=1}^{n-1}\lim_{\varepsilon_0\to 0}\frac{|\varepsilon_j|}{|\varepsilon_0|^{2^{j-1}}}$$
$$= \frac{|f'(\alpha)\rho^{(n-1)}(\alpha)|}{(n-1)!}\prod_{j=1}^{n-1}A_j. \qquad (17)$$

The expression for $\rho^{(n-1)}(\alpha)$ in terms of f can be obtained by successive n times differentiation of $f(t)\rho(t) = t - \alpha$ at the point $t = \alpha$. Using Leibnitz rule, this leads to a system of equations written in matrix form:

$$\begin{bmatrix} \binom{1}{1}f'(\alpha) & 0 & 0 & \cdots & 0 \\ \binom{2}{2}f''(\alpha) & \binom{2}{1}f'(\alpha) & 0 & \cdots & 0 \\ \binom{3}{3}f'''(\alpha) & \binom{3}{2}f''(\alpha) & \binom{3}{1}f'(\alpha) & \cdots & 0 \\ & & \cdots & & \\ \binom{n}{n}f^{(n)}(\alpha) & \binom{n}{n-1}f^{(n-1)}(\alpha) & \binom{n}{n-2}f^{(n-2)}(\alpha) & \cdots & \binom{n}{1}f'(\alpha) \end{bmatrix} \cdot \begin{bmatrix} \rho(\alpha) \\ \rho'(\alpha) \\ \rho''(\alpha) \\ \vdots \\ \rho^{(n-1)}(\alpha) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Cramer's formula for $\rho^{(n-1)}(\alpha)$ reads

$$\rho^{(n-1)}(\alpha) = \frac{(-1)^{n+1}}{n!(f'(\alpha))^n}\begin{vmatrix} \binom{2}{2}f''(\alpha) & \binom{2}{1}f'(\alpha) & 0 & \cdots & 0 \\ \binom{3}{3}f'''(\alpha) & \binom{3}{2}f''(\alpha) & \binom{3}{1}f'(\alpha) & \cdots & 0 \\ & & \cdots & & \\ \binom{n-1}{n-1}f^{(n-1)}(\alpha) & \binom{n-1}{n-2}f^{(n-2)}(\alpha) & \binom{n-1}{n-3}f^{(n-3)}(\alpha) & \cdots & \binom{n-1}{1}f'(\alpha) \\ \binom{n}{n}f^{(n)}(\alpha) & \binom{n}{n-1}f^{(n-1)}(\alpha) & \binom{n}{n-2}f^{(n-2)}(\alpha) & \cdots & \binom{n}{2}f''(\alpha) \end{vmatrix}. \quad \square \qquad (18)$$

Remark 4.1. From the proof of Theorem 4.1 we note that it holds not only for *NS*-iteration, but for every optimal iteration scheme using function sampling cases 1° or 2°, defined in Section 2.

Divided differences are a generalization of derivatives, therefore Theorem 4.1 provides a multipoint version statement of König's result [24]:

Theorem 4.2 (*König*)**.** *Let*

$$g(z) = c_0(x) + c_1(x)(z - x) + c_2(x)(z - x)^2 + \cdots$$

be an analytic function in the disk $|z - x| < \delta$ *centred at* x *and* g *has a single pole at the point* α *belonging to this disk. If* $|\alpha - x| < \sigma\delta < \delta$, *then*

$$\frac{c_k(x)}{c_{k+1}(x)} = \alpha - x + \mathcal{O}(\sigma^{k+2}).$$

Corollary 4.1. *Let* $g(z)$ *be an analytic function in the disk* $D : |z - x| < \delta$, *and* g *has a single simple pole at the point* α *belonging to* D. *Let* $y_j = \Psi_j(x) \in D$ *for* $x \in D$, $j = 0, 1, \ldots, n - 1$, *and* $\varepsilon_j = \alpha - y_j$. *Then*

$$\frac{g[y_0, \ldots, y_{n-2}]}{g[y_0, \ldots, y_{n-1}]} = \alpha - y_{n-1} + \mathcal{O}\Big(\prod_{j=0}^{n-1}\varepsilon_j\Big).$$

The strength of König's type theorems is evident from their application. Let α be a simple isolated zero of an analytic function $f \in D_f$ and $h(z)$ an analytic function in a neighbourhood of α such that $h(\alpha) \neq 0$. Then $g(z) = \frac{h(z)}{f(z)}$ is a meromorphic function from Corollary 4.1. The existence of the function $h(z)$ provides a possibility of choice and results in a family of iterative procedures (17) for the solution of a nonlinear scalar equation $f(x) = 0$. The use of the function $h(z)$ can also be considered as a mean to simplify f evaluation, where applicable.

5. Rational family

We propose a new iterative scheme, with or without derivatives, for the solution of nonlinear equation

$$\begin{cases} y_0 = x_k, \qquad y_1 = y_0 + \gamma f(y_0), \\ y_j = y_{j-1} + \dfrac{\frac{1}{f}[y_{j-2}, \dots, y_0]}{\frac{1}{f}[y_{j-1}, \dots, y_0]}, \quad j = 2, 3, \dots, n, \\ x_{k+1} = y_n. \end{cases} \tag{19}$$

As pointed in Remark 3.1, Newton's and Steffensen's methods are both of the form (8) used in (19). This allows a compact formulation of the family (19). Using a traditional form of Newton's and Steffensen's methods, the family (19) reads

$$\begin{cases} y_0 = x_k, \quad y_1 = y_0 + \gamma f(y_0), \\ y_2 = y_0 - \dfrac{f(y_0)}{f[y_0, y_1]} \\ y_j = y_{j-1} + \dfrac{\frac{1}{f}[y_{j-2}, \dots, y_0]}{\frac{1}{f}[y_{j-1}, \dots, y_0]}, \quad j = 3, \dots, n, \\ x_{k+1} = y_n. \end{cases}$$

Theorem 5.1. *Iteration procedure (19) is an optimal NS-iteration.*

Proof. Using induction and Theorem 4.1 the optimality holds. We can derive the specific asymptotic error constant for method (19).

Obviously, $A_1 = |1 + \gamma f'(\alpha)|$. From (17) for $j \geq 2$ we obtain

$$A_j = \frac{|f'(\alpha)\rho^{(j-1)}(\alpha)|}{(j-1)!} \prod_{k=1}^{j-1} A_k.$$

By induction we proceed to

$$A_j = |(1 + \gamma f'(\alpha))f'(\alpha)|^{2^{j-2}} \frac{|\rho^{(j-1)}(\alpha)|}{(j-1)!} \prod_{k=1}^{j-2} \left(\frac{\rho^{(k)}(\alpha)}{k!} \right)^{2^{j-k-2}}.$$

Expressions for $\rho^{(k)}(\alpha)$ from (18) can further be substituted in A_j. □

Defined in this manner the new approximation in each step can be obtained using only $\mathcal{O}(j)$ rational operations based on the divided differences table in addition to j function evaluations.

An equivalent form of iteration (19) can be observed from the following equality

$$\prod_{i=2}^{j} (y_i - y_{i-1}) = \frac{\frac{1}{f}[y_0]}{\frac{1}{f}[y_{j-1}, \dots, y_0]}.$$

Particular examples of the family (10) or (19) of order four, eight and sixteen were designed (not necessarily with the same approach) in earlier investigations. We cite here some of them.

Derivative free methods:

The scheme (19) of order 4 and $n = 3$

$$[25]: \begin{cases} w_k = x_k + \gamma f(x_k), \qquad y_k = x_k - \dfrac{f(x_k)}{f[x_k, w_k]}, \\ x_{k+1} = y_k - \dfrac{f(x_k)f(w_k)}{f(x_k)f(w_k) - f(x_k)f(y_k) - f(y_k)f(w_k)} \cdot \dfrac{f(y_k)}{f[x_k, w_k]}. \end{cases}$$

The scheme (19) of order 8 and $n = 4$

$$[26]: \begin{cases} w_k = x_k + \gamma f(x_k), \qquad y_k = x_k - \dfrac{f(x_k)}{f[x_k, w_k]}, \\ z_k = y_k - \dfrac{f(x_k)f(w_k)}{f(x_k)f(w_k) - f(x_k)f(y_k) - f(y_k)f(w_k)} \cdot \dfrac{f(y_k)}{f[x_k, w_k]}, \\ x_{k+1} = z_k - \dfrac{(P-Q)(z_k - w_k)(z_k - x_k)}{P(z_k - x_k) + Q(x_k + w_k - y_k - z_k) + R(y_k - w_k)}, \\ P = f[x_k, w_k]f(y_k)f(z_k), \qquad Q = f[y_k, x_k]f(z_k)f(w_k), \\ R = f[z_k, y_k]f(x_k)f(w_k). \end{cases}$$

Newton-like methods:
The scheme (19) of order 4 and $n = 3$

$$\text{Ostrowski's method [5]}: \begin{cases} y_k = x_k - \dfrac{f(x_k)}{f'(x_k)}, \\ x_{k+1} = y_k - \dfrac{f(x_k)}{f(x_k) - 2f(y_k)} \cdot \dfrac{f(y_k)}{f'(x_k)}. \end{cases}$$

The scheme (10) of order 8 and $n = 4$

$$[6]: \begin{cases} y_k = x_k - \dfrac{f(x_k)}{f'(x_k)}, \\ z_k = y_k - \dfrac{f(x_k) + \beta f(y_k)}{f(x_k) + (\beta - 2)f(y_k)} \cdot \dfrac{f(y_k)}{f'(x_k)}, \\ x_{k+1} = x_k - \dfrac{P + Q + R}{Pf[z_k, x_k] + Qf'(x_k) + Rf[y_k, x_k]} f(x_k), \\ P = (x_k - y_k)f(x_k)f(y_k), \qquad Q = (y_k - z_k)f(y_k)f(z_k), \\ R = (z_k - x_k)f(z_k)f(x_k). \end{cases}$$

Acknowledgements

Authors would like to thank the Reviewers for their valuable comments.

References

[1] J.F. Traub, Iterative Methods for the Solution of Equations, Prentice-Hall, Englewood Cliffs, New Jersey, 1964.
[2] M.S. Petković, B. Neta, L.D. Petković, J. Džunić, Multipoint Methods for Solving Nonlinear Equations, Elsevier, Amsterdam, Boston, Heidelberg, 2013.
[3] A. Iliev, N. Kyurkchiev, Nontrivial Methods in Numerical Analysis: Selected Topics in Numerical Analysis, LAP LAMBERT Academic Publishing, Saarbrücken, 2010.
[4] M.S. Petković, B. Neta, L.D. Petković, J. Džunić, Multipoint methods for solving nonlinear equations: a survey, Appl. Math. Comput. 226 (2014) 635–660.
[5] A.M. Ostrowski, Solution of Equations and Systems of Equations, Academic Press, New York, 1960.
[6] J.R. Sharma, R.K. Guha, P. Gupta, Improved Kings methods with optimal order of convergence based on rational approximations, Appl. Math. Lett. 26 (2013) 473–480.
[7] R. Behl, S. Amat, S.S. Montsa, A new highly efficient and optimal family of eighth-order methods for solving nonlinear equations, Appl. Math. Comput. 282 (2016) 175–186.
[8] L. Tornheim, Convergence of multipoint iterative methods, J. ACM 11 (1964) 210–220.
[9] P. Jarratt, D. Nudds, The use of rational functions in the iterative solution of equations on a digital computer, Comput. J. 8 (1965) 62–65.
[10] F.M. Larkin, Root-finding by fitting rational functions, Math. Comp. 35 (1980) 803–816.
[11] G.A. Baker, P.G. Morris, Padé Approximants, Cambridge University Press, 2009, (online).
[12] O.P. Bruno, F. Reitich, Approximation of Analytic Functions: a Method of Enhanced Convergence, Research Report No 92-NA-021, Carnegie Mellon University, 1992.
[13] P. Kowalczyk, On root finding algorithms for complex functions with branch cuts, J. Comput. Appl. Math. 314 (2017) 1–9.
[14] H.T. Kung, J.F. Traub, Optimal order of one-point and multipoint iteration, J. ACM 21 (1974) 643–651.
[15] G.M. Phillips, Interpolation and Approximation by Polynomials, Springer-Verlag, New York Inc., 2003.
[16] L.M. Milne-Thomson, The Calculus of Finite Differences, McMillan and Co., London, 1933.
[17] T. Popoviciu, Sur quelques propriétés des fonctions dune ou de deux variables reélles (dissertation), at Faculté des Sciences de Paris, 1933.
[18] T. Popoviciu, Introduction á la théorie des différences divisées, Bull. Math. Soc. Roumanie Sci. 42 (1940) 65–78.
[19] J.F. Steffensen, Note on divided differences, Danske Vid. Selskab. Mat.-Fys. Medd. 17 (1939) 1–12.
[20] J.F. Steffensen, Remarks on iteration, Skand. Aktuarietidskr. 16 (1933) 64–72.
[21] M.C. Schneider, K.W. Werner, Hermite interpolation: The Barycentric approach, Computing 46 (1991) 35–51.
[22] J. Berrut, G. Klein, Recent advances in linear barycentric rational interpolation, J. Comput. Appl. Math. 259 (2014) 95–107.
[23] M.S. Floater, T. Lyche, Two chain rules for divided differences and Faa di Brunos formula, Math. Comp. 258 (2007) 867–877.
[24] J. König, Über eine Eigenschaft der Potenzreihen, Math. Ann. 23 (1884) 447–449.
[25] M.S. Petković, S. Ilić, J. Džunić, Derivative free two-point methods with and without memory for solving nonlinear equations, Appl. Math. Comput. 217 (2010) 1887–1895.
[26] X. Wang, J. Džunic, T. Zhang, On an efficient family of derivative free three-point methods for solving nonlinear equations, Appl. Math. Comput. 219 (2012) 1749–1760.