On an Application of Herzberger’s Matrix Method to Multipoint Families of Root-solvers

Jovana Džunić, Ljiljana D. Petković, Miodrag S. Petković

Faculty of Electronic Engineering, University of Niš, Serbia
Faculty of Mechanical Engineering, University of Niš, Serbia

Abstract. An application of Herzberger’s matrix method, very rarely used in the topic of multipoint methods for solving nonlinear equations, is presented. It is shown that the area of application of Herzberger’s matrix method is wider than it is presented in [J. Herzberger, Über Matrixdarstellungen für Iterationverfahren bei nichtlinearen Gleichungen, Computing, 12 (1974) 215–222]. This method is applied for the determination of the order of convergence of multipoint families of methods, Steffensen’s type and Newton’s type, with and without memory. The advantage and the elegance of this method arise from ease in handling matrices.

In Memory of Professor Jürgen Herzberger.

1. Introduction

In this paper we demonstrate the application of Herzberger’s matrix method [1] (HMM for short) for calculation of the order of convergence of multipoint methods for solving nonlinear equations. Herzberger has presented a matrix procedure which reduces the problem of determination of the order of convergence of Hermite’s class of one-point methods, single-step and total-step methods to the problem of finding the spectral radius of a certain matrix. We prove that Herzberger’s approach can be used to an arbitrary class of multipoint methods with or without memory. The requirement is the knowledge of an error relation of the method considered.

The main advantage of multipoint methods is their high computational efficiency since they can attain convergence order $2^n$ consuming only $n + 1$ function evaluations. Methods with this property are usually called optimal methods and support the Kung-Traub hypothesis [2] on the upper bound of convergence order of multipoint methods. For example, Kung-Traub’s families [2], Zheng-Li-Huang’s family [3], a general class based on Hermite’s interpolation polynomials [4] and on Hermite’s interpolation by rational functions [5], belong to the class of $n$–point optimal methods of general form. In our study we shall consider both Newton’s type and Steffensen’s type methods with and without memory. It will be shown that under the
same amount of Hermite’s sampled data, Steffensen’s type of methods produce higher convergence order for methods with memory.

The paper is organized as follows. In Section 2 we recall some basic definitions and theorems of matrix and Perron-Frobenius theory. Section 3 is devoted to Herzberger’s work [1] and provides evidence of applicability of HMM in a broader sense than that defined in [1]. In the next section application of HMM is explored. A multipoint derivative free class of iterative methods without memory, which starts with an arbitrary two-point method of fourth order, has been proposed in a condensed form as a five-line note in [6]. In Section 4 we present this class of methods in an expanded form and its generalization, suitable for the application of HMM. With HMM we give an alternative proof that the order of these two multipoint families is \(2^n\). In Section 4 we also use HMM to compare acceleration effects using information from the current and one previous iteration to Newton’s and Steffensen’s type of multipoint methods.

2. Non-negative Matrices

We recall some basic notions of matrix theory regarding positive and non-negative matrices.

**Definition 2.1.** A real \(n \times m\) matrix \(A = [a_{ij}]\) is non-negative if

\[ a_{ij} \geq 0, \quad \forall i, j, \quad i \in \{1, 2, \ldots, n\}, \quad j \in \{1, 2, \ldots, m\}. \]

If \(a_{ij} > 0, \forall i, j, \quad i \in \{1, 2, \ldots, n\}, \quad j \in \{1, 2, \ldots, m\}\), then the matrix \(A\) is positive.

**Definition 2.2.** A non-negative \(n \times n\) matrix \(A\) is primitive if there exists \(k \in \mathbb{N}\) such that \(A^k\) is positive.

**Definition 2.3.** A permutation matrix is a square binary matrix that has exactly one entry of 1 in each row and each column, and zeros elsewhere.

**Definition 2.4.** A matrix is irreducible if it is not similar via a permutation to a block upper triangular matrix, that has more than one block of positive size.

For a reducible matrix \(A\) there is a permutation matrix \(P\) such that

\[ PAP^T = \begin{bmatrix} X & Y \\ O & Z \end{bmatrix}, \]

where \(X\) and \(Z\) are square matrices of order at least 1. \(O\) denotes a rectangular zero-matrix. This definition is equivalent to the existence of a permutation matrix \(Q\) such that

\[ QAQ^T = \begin{bmatrix} Z & O \\ Y & X \end{bmatrix}. \]

A list of theorems used in the rest of the paper and appropriate references are given as follows. These results originate from [7] and [8] and they can be found, e.g., in [9] or [10].

**Theorem 2.5.** Every primitive matrix is irreducible.

**Theorem 2.6 (Perron-Frobenius).** Let \(A\) be a non-negative irreducible \(n \times n\) matrix. Then,

1. \(A\) has a positive real eigenvalue equal to its spectral radius \(\rho(A)\).
2. To \(\rho(A)\) there corresponds a positive eigenvector \(x\).
3. \(\rho(A)\) increases when any entry of \(A\) increases.
4. \(\rho(A)\) is a simple eigenvalue of \(A\).
Theorem 2.7. Let $A$ be a non-negative irreducible $n \times n$ matrix. There exists an $n \times n$ matrix $B$ such that
\[
\lim_{j \to \infty} \left( \frac{A}{\rho(A)} \right)^j = B
\]
if and only if $A$ is primitive.

Theorem 2.8. If $A = [a_{ij}]$ is a non-negative irreducible $n \times n$ matrix, then either
\[
\sum_{j=1}^{n} a_{ij} = \rho(A) \quad \text{for all} \quad 1 \leq i \leq n,
\]
or
\[
\min_{1 \leq i \leq n} \sum_{j=1}^{n} a_{ij} < \rho(A) < \max_{1 \leq i \leq n} \sum_{j=1}^{n} a_{ij}.
\]

For matrix manipulations we use Sylvester’s theorem stated in 1851 (see [11] for details and variants of Sylvester’s theorem).

Theorem 2.9 (Sylvester). If $A$ and $B$ are matrices of size $m \times k$ and $k \times m$ ($m \geq k$), respectively, then
\[
\det(\lambda I_m - AB) = \lambda^{m-k} \det(\lambda I_k - BA),
\]
where $I_j$ is the identity matrix of order $j$.

3. Herzberger’s Matrix Method

Iterative root-finding methods are one of the oldest and most persistent subjects of investigation. Basic problem is to approximate a zero $\alpha$ of a function $f$ by iterative means. This has evolved to finding an iteration function that would solve such a problem in the least of time with the prescribed amount of data. This is in close relation to order of convergence and optimality of an iterative function. In [1] a connection was established between order of convergence of an iterative method and eigenvalues of a matrix formed in a special manner. The connection was determined through an error relation of the analysed iteration. In [1] Herzberger organized presentation around one-point interpolatory iteration functions $\varphi$ with memory of the form
\[
x_{k+1} = \varphi(x_k, x_{k-1}, \ldots, x_{k-n}),
\]
due to the beforehand knowledge of their error relation. The specific estimate reads
\[
|x_{k+1} - \alpha| = C_k \prod_{j=0}^{n} |x_{k-j} - \alpha|^{m_j},
\]
where $m_j \in \mathbb{N}$ is the number of function evaluations calculated at the point $x_{k-j}$: $f(x_{k-j}), f'(x_{k-j}), \ldots, f^{(m_j-1)}(x_{k-j})$ and for some expression
\[
C_k = C(x_k, x_{k-1}, \ldots, x_{k-n}) \to C = \text{const. as } x_k \to \alpha.
\]

This class of one-point interpolatory root-finding algorithms was coined Hermite interpolatory iteration functions (HIF for short) because it operates on Hermite interpolation (both direct and indirect).

We wish to prove that HMM has a wider range of use to that defined in [1]. Specifically, HMM is applicable for stationary iterative methods whenever the error relation of an iterative method is known and takes form (1) for $m_j \geq 0, j \in \{0, 1, \ldots, n\}$. Therefore, we are broadening the domain of coefficients $m_j$ from $\mathbb{N}$ to $\mathbb{N}_0$ and the application of HMM to a wider class of methods. This will particularly become of importance when analysing multipoint iterations.

A wider application range of Herzberger’s method (HMM) is obtained by adapting the original proof from [1] to validate conclusions for a broader class of single root finding methods.
Theorem 3.1. Let the error relation (1) be valid for some iterative root finding method \( \varphi = \varphi(x_k, x_{k-1}, \ldots, x_{k-n}) \), where all \( m_j \geq 0 \). Then the order of convergence of the stationary iteration

\[ x_{k+1} = \varphi(x_k, x_{k-1}, \ldots, x_{k-n}), \quad k \geq n, \quad k \in \mathbb{N}, \]

equals the spectral radius of the associated matrix

\[
M(\varphi) = \begin{bmatrix} m_0 & m_1 & m_2 & \cdots & m_{n-1} & m_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.
\] (2)

Proof. Having in mind that we are analyzing error relation of an iteration function, we will not consider the trivial case when all \( m_j = 0 \).

First, it is assumed that all \( m_j > 0 \). Introduce

\[ \epsilon_k = -\log |x_k - \alpha|, \quad c_k = -\log C_k, \]

and apply logarithm to (1). The following is obtained

\[ \epsilon_{k+1} = c_k + \sum_{j=0}^{n} m_j \epsilon_{k-j}. \] (3)

Convergence of \( \{x_k\} \) to \( \alpha \) is assumed and that the order of convergence exists and equals

\[ r = \lim_{j \to \infty} \frac{\epsilon_{j+1}}{\epsilon_j}. \]

Divide (3) by \( \epsilon_{k-n} \) and allow \( k \to \infty \). Since \( \epsilon_k \to \infty \) and

\[ \lim_{j \to \infty} \frac{\epsilon_{j+1}}{\epsilon_j} = \lim_{j \to \infty} \frac{\epsilon_{j+1}}{\epsilon_{j+1}} \frac{\epsilon_{j+1}}{\epsilon_{j+2}} \cdots \frac{\epsilon_{j+1}}{\epsilon_j} = r', \]

we obtain that the order of convergence \( r \) satisfies the equation

\[ P_{\varphi}(x) := x^{n+1} - \sum_{j=0}^{n} m_j x^{n-j} = 0. \] (4)

By Descartes’ rule of signs, \( P_{\varphi} \) defined in (4) has a unique positive root - the order of convergence \( r \). Note that Descartes’ rule of signs provides the same conclusion in the case when some \( m_j = 0 \). The Frobenius companion matrix [12] of the polynomial (4) reads

\[
M(\varphi) = \begin{bmatrix} m_0 & m_1 & \cdots & m_{n-1} & m_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}.
\] (5)

Characteristic polynomial and the minimal polynomial of \( M(\varphi) \) coincide with \( P_{\varphi}(\lambda) = \det \left( M_{n+1} - M(\varphi) \right) \), and \( M(\varphi) \) is a primitive matrix (see e.g. [9], [10], [12]), thus \( r = \rho(M(\varphi)) \).

Now we shall consider the case when some \( m_j \) vanish. There exists an array of matrices \( \{ M_k \}_{k \in \mathbb{N}} \) of the form (5) with all positive entries in the first raw that converge to the matrix \( M(\varphi) \). Basically, if \( m_j = 0 \) in \( M(\varphi) \) we substitute it with a positive value \( \delta_k \) to form the matrix \( M_k \), and then allow \( \delta_k \) to be arbitrarily close to zero as \( k \) approaches infinity. Each \( M_k \) has a unique positive eigenvalue \( r_k = \rho(M_k) \). Polynomial zeros are continuous functions of its coefficients, so by the argument of continuity we can conclude that the spectral radius of \( M(\varphi) \) is its eigenvalue. Therefore, the order of convergence \( r \) of the iteration function \( \varphi \) equals spectral radius of the matrix \( M(\varphi) \). It is also the unique positive eigenvalue of multiplicity one, by Descartes’ rule of signs. □
Remark 3.1. The error relation (3) can as well be presented in a matrix form

\[
\begin{bmatrix}
\epsilon_{k+1} \\
\epsilon_k \\
\epsilon_{k-1} \\
\vdots \\
\epsilon_{k-n+1}
\end{bmatrix} =
\begin{bmatrix}
m_0 & m_1 & \cdots & m_{n-1} & m_n \\
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 1 & 0
\end{bmatrix}
\begin{bmatrix}
\epsilon_k \\
\epsilon_{k-1} \\
\epsilon_{k-2} \\
\vdots \\
\epsilon_{k-n}
\end{bmatrix} +
\begin{bmatrix}
0 \\
0 \\
0 \\
\vdots \\
0
\end{bmatrix},
\]

which establishes an even more profound and natural connection between iteration formula \( \varphi \) and its associated matrix \( M(\varphi) \).

Remark 3.2. Iterative root finding methods are applied in real life problems only when their order of convergence \( r \geq 1 \). This is in accordance with a demand that there is no sign change of \( P_\varphi(x) \) between 0 and 1, or equivalently
\[
P_\varphi(1) \leq 0 \iff m_0 + \cdots + m_n \geq 1.
\]

For \( m_j \in \mathbb{N}_0 \) where not all \( m_j \) vanish, this inequality holds.

Consideration of the case when some \( m_j \) take zero value is important for analysis of iterations without memory and multipoint iterations. For example, the matrix \( M(N) \), associated to Newton’s method \( N \), takes one of the forms
\[
M(N) = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix}, \quad M(N) = \begin{bmatrix} 2 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \ldots
\]

This is a consequence of the fact that any iteration can be considered as a function of more variables where some of these variables do not participate in the error relation. Not all matrices associated to Newton’s method are primitive or irreducible, however HMM formula works as much expected.

With powerful numerical tools for estimating dominant eigenvalues and modern computers, HMM method allows ease in computing order of convergence of one point iterations with memory. Newton’s method can be successfully used to determine the dominant eigenvalue in this case. Results such as those presented in [13] and [14] provide good initial value estimates.

The application of HMM expands to compositions of iteration functions. Composition of functions is naturally related to matrix multiplication. Using Perron-Frobenius theory of irreducible matrices it was proved in [1] that the order of convergence of a composition of HIF iterations \( \varphi = \varphi_m \circ \varphi_{m-1} \circ \cdots \circ \varphi_1 \) with associated matrices \( M_m, M_{m-1}, \ldots, M_1 \), where \( M_k = M(\varphi_k) \), preserves feature:
\[
r(\varphi) = \rho(M_m \cdot M_{m-1} \cdots M_1), \quad \text{and} \quad r(\varphi) \in Sp(M_m \cdot M_{m-1} \cdots M_1),
\]

where \( Sp(M) \) is the spectrum of the matrix \( M \). We shall prove that this feature is not restricted to interpolatory type of iteration functions. Also, matrices \( M_j \) need not be primitive while of the specific form.

Composition of one-point iteration functions with memory \( \varphi_1 \) and \( \varphi_2 \) is defined through extensions \( \varphi_j : D^{n+1} \to D^{n+1}, \ j = 1, 2 \)
\[
\varphi_j(x_k, x_{k-1}, \ldots, x_{k-n}) = (\varphi_j(x_k, x_{k-1}, \ldots, x_{k-n}), x_k, \ldots, x_{k-n+1})
\]

where \( D \subset \mathbb{C} \) is some open set containing \( a \). Let \( y_{k+1} = \varphi_1(x_k, x_{k-1}, \ldots, x_{k-n}) \). Composition \( \varphi = \varphi_1 \circ \varphi_2 \) is defined by an implicit equation
\[
(\varphi(x_k, x_{k-1}, \ldots, x_{k-n}), y_{k+1}, x_k, \ldots, x_{k-n+2}) = (\varphi_2 \circ \varphi_1)(x_k, x_{k-1}, \ldots, x_{k-n}).
\]

In other words,
\[
\varphi(x_k, x_{k-1}, \ldots, x_{k-n}) = \varphi_2(\varphi_1(x_k, x_{k-1}, \ldots, x_{k-n}), x_k, \ldots, x_{k-n+1}).
\]
**Theorem 3.2.** Let iterations \( q_j = q_j(x_k, x_{k-1}, \ldots, x_{k-n}) \) all have error relations

\[
|x_{k+1} - \alpha| = C^{(i)}_k \prod_{i=0}^n |x_{k-i} - \alpha|^{m_i^{(i)}},
\]

where all \( m_i^{(i)} \in \mathbb{N}_0 \). Then the order of convergence of the stationary iteration

\[
x_{k+1} = q(x_k, x_{k-1}, \ldots, x_{k-n}) = (q_m \circ q_{m-1} \circ \cdots \circ q_1)(x_k, x_{k-1}, \ldots, x_{k-n})
\]

equals the spectral radius of the matrix

\[
M = M_m \cdot M_{m-1} \cdots M_1, \quad M_j = M(q_j).
\]

**Proof.** We first consider the case when all \( m_i^{(i)} > 0 \). It can be verified that any product of associated matrices \( M_j = M(q_j) \) with positive first rows is still a primitive matrix. Also, for the entries \( a_{ij} \) of the product of \( M_i \) it is easily verified that they are all nonnegative integers and \( \min_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} \geq 1 \) and \( \max_{1 \leq i \leq n} \sum_{j=1}^n a_{ij} > 1 \). From Theorem 2.8 it follows \( \rho(\prod M_i) > 1 \).

Error relations (7) for iterations \( q_j, \ j = 1, 2, \ldots, m, \) in matrix form (6) read

\[
\begin{bmatrix}
\vdots \\
c_{k+1} \\
e_k \\
c_{k-1} \\
\vdots \\
c_{k-n+1}
\end{bmatrix}
= \begin{bmatrix}
m_0^{(i)} \\
m_1^{(i)} \\
m_2^{(i)} \\
m_n^{(i)}
\end{bmatrix}
\begin{bmatrix}
c_{k+1} \\
e_k \\
c_{k-1} \\
\vdots \\
c_{k-n+1}
\end{bmatrix}
= \begin{bmatrix}
\cdots \\
0 \\
1 \\
0 \\
\cdots
\end{bmatrix}
\begin{bmatrix}
\cdots \\
0 \\
1 \\
0 \\
\cdots
\end{bmatrix}
\begin{bmatrix}
c_{k+1} \\
e_k \\
c_{k-1} \\
\vdots \\
c_{k-n+1}
\end{bmatrix}
= \prod_{j=1}^n (\prod_{i=j+1}^m M_i) \cdot c_j^{(i)},
\]

where \( c_j^{(i)} \to c^{(i)} \) when \( k \to \infty \). For iteration \( q \) we thus have

\[
e_{k+1} = M_m \cdot M_{m-1} \cdots M_1 \cdot e_k + v_k = M \cdot e_k + v_k,
\]

where

\[
e_k = \begin{bmatrix}
e_k \\
e_{k-1} \\
\vdots \\
e_{k-n+1}
\end{bmatrix}^T, \quad M = M_m \cdot M_{m-1} \cdots M_1,
\]

\[
v_k = \sum_{j=1}^m \left( \prod_{i=j+1}^m M_i \right) \cdot c_j^{(i)},
\]

\[
c_j^{(i)} = \begin{bmatrix}
c_{j+1}^{(i)} \\
0 \\
\vdots \\
0
\end{bmatrix}^T.
\]

Note that components of the vector \( v_k \) converge to some constant expression \( v \) as \( k \to \infty \), that is \( v_k \to v = [v \ v \ \cdots \ v]^T \), when \( k \to \infty \). Since every convergent sequence is bounded, there exists \( \delta > 0 \) such that

\[
||v_k - v|| < \delta, \text{ for all } k \in \mathbb{N}.
\]

For arbitrary \( j \in \mathbb{N} \) we have

\[
e_{k+j} = M^j e_k + \sum_{s=0}^{j-1} M^s v_{k+j-s-1}
= \rho(M)^j \left( \frac{M}{\rho(M)} \right)^j e_k + \sum_{s=0}^{j-1} \rho(M)^s \left( \frac{M}{\rho(M)} \right)^s v_{k+j-s-1}.
\]

\[
\rho(M)\left( \frac{M}{\rho(M)} \right)^j e_k + \sum_{s=0}^{j-1} \rho(M)^s \left( \frac{M}{\rho(M)} \right)^s v_{k+j-s-1}.
\]

\[
J. Džunić et al. / Filomat 32:11 (2018), 3815–3829
\]

3820
Let $B$ be a matrix from Theorem 2.7, such that $\lim_{j \to \infty} \left( \frac{M}{\rho(M)} \right)^j = B$. Sequences

\[
\left\{ \left\| \left( \frac{M}{\rho(M)} \right)^j \right\| \right\}_{j \in \mathbb{N}} \quad \text{and} \quad \left\{ \left\| \left( \frac{M}{\rho(M)} \right)^j - B \right\| \right\}_{j \in \mathbb{N}}
\]

are convergent and are bounded. There exist $\delta_1, \delta_2 > 0$ such that for all $j \in \mathbb{N}$

\[
\left\| \left( \frac{M}{\rho(M)} \right)^j \right\| < \delta_1, \quad \left\| \left( \frac{M}{\rho(M)} \right)^j - B \right\| < \delta_2.
\]

Error relation (11) may be written in the form,

\[
e_{k+1} = \rho(M)^j e_k + \sum_{s=0}^{j-1} \rho(M)^s \left( \frac{M}{\rho(M)} \right)^{j-s-1} v + \left( \frac{M}{\rho(M)} \right)^j B v + B v.
\]

Using properties of the induced matrix norm, vector norm and modulus (triangle inequalities)

\[
\left\| \left( \frac{M}{\rho(M)} \right)^j \right\| \leq \left\| a \right\| \leq \left\| a + b \right\| \leq \left\| a \right\| + \left\| b \right\|,
\]

from (10)–(13) it follows

\[
\left\| e_{k+1} \right\| \leq \rho(M)^j \left\| e_k \right\| + \varepsilon \sum_{s=0}^{j-1} \rho(M)^s
\]

\[
= \rho(M)^j \left\| e_k \right\| + \varepsilon \frac{\rho(M)^j - 1}{\rho(M) - 1}
\]

\[
\leq \rho(M)^j \left( \left\| e_k \right\| + \frac{\varepsilon}{\rho(M) - 1} \right),
\]

and

\[
\left\| e_{k+1} \right\| \geq \rho(M)^j \left( \left\| e_k \right\| \right) - \frac{\varepsilon}{\rho(M) - 1},
\]

for all $j \in \mathbb{N}$, and $\varepsilon = \delta_1 \delta + \delta_2 \|v\| + \|B v\|$. Having in mind $\left\| \left( \frac{M}{\rho(M)} \right)^j e_k \right\| \to \infty$, $k \to \infty$, for $k \in \mathbb{N}$ large enough the inequality

\[
\left\| \left( \frac{M}{\rho(M)} \right)^j e_k \right\| - \frac{\varepsilon}{\rho(M) - 1} > 0
\]

is valid.

We conclude

\[
0 < \rho(M)^j \left( \left\| e_k \right\| + \alpha_k \right) \leq \left\| e_{k+1} \right\| \leq \rho(M)^j \left( \left\| e_k \right\| + \beta_k \right),
\]

where for fixed $k$, $\alpha_k$ and $\beta_k$ are some constants.

Among several equivalent ways to obtain the order of convergence $r$, in [15] (see also [16]) it was proved that $r = \lim_{j \to \infty} \|e_j\|^{1/j}$. Taking $(k + j)$–th root of inequalities (14) and allowing $j \to \infty$, relying on limits:

\[
\lim_{j \to \infty} \left\| \left( \frac{M}{\rho(M)} \right)^j e_k \right\| = \|B e_k\| \quad \text{and} \quad \lim_{j \to \infty} \sqrt[2k]{C} = 1 \quad \text{we get} \quad r = \rho(M).
\]

As in the conclusion of the proof of Theorem 3.1, using the argument of continuity we verify the statement in the case when some $m^{(i)}_i$ are equal to zero. □
When considering multipoint methods without memory, the following is observed: Multipoint methods are a composition \( \varphi = \varphi_m \circ \varphi_{m-1} \circ \cdots \circ \varphi_1 \) of one-point iteration functions with/without memory \( \varphi_2, \varphi_3, \ldots, \varphi_m \) and \( \varphi_1 \) a one-point iteration function without memory. For example, Traub-Steffensen’s method (S)

\[
x_{k+1} = x_k - \frac{\gamma f(x_k)^2}{f(x_k + \gamma f(x_k)) - f(x_k)}, \quad \gamma \neq 0,
\]

can be viewed as a composition of Secant method (\( \varphi_2 \)) and Traub-Steffensen’s correction (\( \varphi_1 \)):

\[
\varphi_2(x_k) = x_k - \frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}} \quad \text{and} \quad \varphi_1(x_k) = x_k + \gamma f(x_k).
\]

The associated matrices read

\[
M(\varphi_2) = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad M(\varphi_1) = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.
\]

Their product, according to (8), gives

\[
M(S) = M_2 \cdot M_1 = \begin{bmatrix} 2 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \rho(M(S)) = 2.
\]

Matrix \( M(\varphi_1) \) is neither a primitive matrix, nor it is irreducible. However, HMM can still be used to calculate the order of convergence.

Additional examples of the application of HMM methods can be found in the early works [18], [19] of B. Neta.

Multipoint methods with memory are a composition of one-point iteration functions with/without memory \( \varphi = \varphi_m \circ \varphi_{m-1} \circ \cdots \circ \varphi_1 \). With such interpretation it is obvious that HMM can be used to calculate the order of convergence of multipoint methods as long as the error relations are known. This application is the subject of the following section. Although in this communication we restrict our attention to the scalar case, it is important to note that HMM is applicable to the non-scalar case as well.

4. Application of Herzberger’s Matrix Method to Multipoint Methods

4.1. Family of multipoint derivative free methods without memory

Let \( \alpha \) be a simple zero of a given function \( f \), and \( x \) an approximation to \( \alpha \). The multipoint family of derivative free methods for solving nonlinear equations proposed in [6]

\[
\begin{align*}
y_{k,0} &= x_k, \quad y_{k,1} = y_{k,0} + \gamma f(y_{k,0}), \quad \gamma \neq 0, \\
y_{k,2} &= y_{k,0} - \frac{f(y_{k,0})}{f(y_{k,1})}, \\
y_{k,3} &= \frac{y_{k,0} f(y_{k,1}) - y_{k,1} f(y_{k,0})}{y_{k,1} - y_{k,0}}, \\
y_{k,j+1} &= y_{k,j} - \frac{f(y_{k,j})}{P_j'(y_{k,j})}, \quad (j = 3, \ldots, n), \\
x_{k+1} &= y_{k,n+1}, \quad k \in \mathbb{N}_0,
\end{align*}
\]

(15)

relies on the method from [3]. It is a generalization due to the fact that pre-conditioner (steps \( y_{k,0} \) to \( y_{k,3} \)) is an arbitrary two-point optimal family of methods starting with Steffensen-Traub’s correction. Steps that define \( y_{k,j+1} \), \( j = 3, \ldots, n \), use Newton’s interpolating polynomials \( P_j(t) = P_j(t; y_{k,0}, \ldots, y_{k,j}) \).
As noted in [6], this approach of increase in convergence by Newton’s polynomial is of general nature. In this manner we can define optimal families of methods

\[
\begin{cases}
y_{k,0} = x_k, \\ y_{k,1} = y_{k,0} + \gamma f(y_{k,0}), \gamma \neq 0, \\ y_{k,2} = y_{k,0} - \frac{f(y_{k,0})}{f[y_{k,0}, y_{k,1}]}, \\ y_{k,j+1} = \theta_2(y_{k,0}, \ldots, y_{k,j}), \quad (j = 2, \ldots, p), \\ y_{k,j+1} = \frac{y_{k,j} - f(y_{k,j})}{p_j(y_{k,j})}, \quad (j = p+1, \ldots, n), \\ x_{k+1} = y_{k,n+1}, \quad k \in \mathbb{N}_0,
\end{cases}
\]

(16)

where \( \theta_2: (y_{k,0}, \ldots, y_{k,j-1}) \) is an arbitrary optimal derivative-free method of order \( 2^{j-1} \). We note that dealing with \( j > 4 \) is only of academic interest, which has been mentioned several times in the existing literature. First we analyze methods (15).

**Theorem 4.1.** The order of convergence of the family of derivative free methods (15) is \( 2^n \).

**Proof.** According to error relations

\[
\begin{align*}
\varepsilon_{k,0} &= y_{k,0} - \alpha, \\
\varepsilon_{k,1} &= y_{k,1} - \alpha \sim \varepsilon_{k,0}, \\
\varepsilon_{k,2} &= y_{k,2} - \alpha \sim \varepsilon_{k,0} \varepsilon_{k,1}, \\
\varepsilon_{k,3} &= y_{k,3} - \alpha \sim \varepsilon_{k,0}^3, \\
\varepsilon_{k,j+1} &= y_{k,j+1} - \alpha \sim \prod_{i=0}^{j} \varepsilon_{k,i}, \quad j = 3, \ldots, n+1, \quad k \in \mathbb{N}_0,
\end{align*}
\]

we form companion matrices of dimensions \((n+1) \times (n+1)\). Indices of companion matrices \( M_j \) match the second index of the approximation \( y_{k,j} \). For brevity, block matrices notation is used, with

\[
\begin{align*}
a_j &= \begin{bmatrix} 1 & 1 & 0 & \ldots & 0 \\ & & & & \\ & & & & \\ 0 & 0 & 0 & \ldots & 0 \\ & & & & \\ n+1-j & & & & \\
\end{bmatrix}, & b &= \begin{bmatrix} 0 & 0 & 4 & 0 & 0 & \ldots & 0 \\ & & & & & & \\ & & & & & & \\ & & & & & & \\
\end{bmatrix}
\]

row matrices, \( O_{jci} = [0]_{jci} \) a zero rectangular matrix and \( I_j \) is the identity matrix of dimension \( j \):

\[
M_j = \begin{bmatrix}
a_j \\
I_n \begin{bmatrix} O_{nx1} \\
\end{bmatrix}
\end{bmatrix}, \quad j \neq 3, \quad M_3 = \begin{bmatrix}
b \\
I_n \begin{bmatrix} O_{nx1} \\
\end{bmatrix}
\end{bmatrix}.
\]

It is easily verified that when \( M_3 \) is not a factor, the following equalities hold

\[
M_{j+1}M_j = \begin{bmatrix}
\frac{2a_j}{a_j} \\
\frac{a_j}{I_{n-1} \begin{bmatrix} O_{(n-1)x2} \end{bmatrix}}
\end{bmatrix}, \quad M_{j+2}M_{j+1}M_j = \begin{bmatrix}
\frac{2^2a_j}{2a_j} \\
\frac{2a_j}{a_j} \\
\frac{a_j}{I_{n-2} \begin{bmatrix} O_{(n-2)x3} \end{bmatrix}}
\end{bmatrix}.
\]

Using induction it is easily demonstrated that when \( M_3 \) is not a factor, the matrix product reads

\[
M_{j+s} \cdots M_j = \begin{bmatrix}
A_{s+1}(j) \\
I_{n-s} \begin{bmatrix} O_{(n-s)x(s+1)} \end{bmatrix}
\end{bmatrix}, \quad 1 \leq s \leq n + 1 - j,
\]

(17)
where \(A_{s+1}(j)\) is the \((s + 1) \times (n + 1)\) matrix with \(i\)-th row equal to \(2^{s+1-i}a_j\), that is

\[
A_{s+1}(j) = \begin{bmatrix}
2^s a_j \\
2^{s-1} a_j \\
\vdots \\
2^0 a_j
\end{bmatrix} = \begin{bmatrix}
j & 2^s & 2^s & 0 & \ldots & 0 \\
2^{s-1} & 2^{s-1} & 2^{s-1} & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1 & 1 & 1 & \ldots & 0
\end{bmatrix}.
\]

(18)

Thus, regard to Theorem 3.2, for the method (15) we obtain

\[
M = (M_{s+1} \cdots M_4)M_3(M_2 : M_1)
\]

\[
= \begin{bmatrix}
A_{n-2}(4) \\
I_3 & O_{3 \times (n-2)}
\end{bmatrix} \cdot M_3 \cdot \begin{bmatrix}
A_2(1) \\
I_{n-1} & O_{(n-1) \times 2}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
A_{n-2}(4) \\
I_3 & O_{3 \times (n-2)}
\end{bmatrix} \cdot \begin{bmatrix}
A_3(1) \\
I_{n-2} & O_{(n-2) \times 3}
\end{bmatrix}
\]

\[
= \begin{bmatrix}
2^{n-3} & 2^{n-3} & 2^{n-3} & 0 & \ldots & 0 \\
2^{n-4} & 2^{n-4} & 2^{n-4} & 0 & \ldots & 0 \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
1 & 1 & 1 & 1 & \ldots & 0 \\
1 & 0 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 & \ldots & 0 \\
0 & 1 & 0 & 0 & \ldots & 0
\end{bmatrix} \cdot \begin{bmatrix}
4 & 0 & \ldots & 0 & 0 & 0 \\
2 & 0 & \ldots & 0 & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\
0 & 1 & \ldots & 0 & 0 & 0 \\
0 & 0 & \ldots & 0 & 0 & 0
\end{bmatrix}
\]

\[
= \begin{bmatrix}
2^n & 0 & \ldots & 0 \\
2^{n-1} & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
1 & 0 & \ldots & 0
\end{bmatrix}.
\]

We now easily conclude \(r = \rho(M) = 2^n\). \(\square\)

**Theorem 4.2.** The order of convergence of the family of derivative free methods (16) is \(2^n\).

**Proof.** Error relations of (16) for \(k \in \mathbb{N}_0\) read

\[
\varepsilon_{k,0} = y_{k,0} - \alpha,
\]

\[
\varepsilon_{k,1} = y_{k,1} - \alpha \sim \varepsilon_{k,0},
\]

\[
\varepsilon_{k,2} = y_{k,2} - \alpha \sim \varepsilon_{k,0} \varepsilon_{k,1} \sim \varepsilon_{k,0}^2,
\]

and

\[
\varepsilon_{k,j} = y_{k,j} - \alpha \sim \begin{cases} 
\varepsilon_{k,0}^{2^{j-1}} & j = 3, \ldots, p+1, \\
\prod_{i=0}^{j-1} \varepsilon_{k,i} & j = p+2, \ldots, n+1.
\end{cases}
\]
Using earlier notation \( a_j = [1 \ldots 1 0 \ldots 0] \) and \( b_j = 2^{j-1}(a_j - a_{j-1}) \), \( b_1 = a_1 \), \( O_{p|x} = [0]_{p|x} \) a zero rectangular matrix and \( I_j \) the identity matrix of dimension \( j \), associated matrices are

\[
M_j = \begin{bmatrix}
    b_j & I_n \mid O_{nx1} \\
    a_j & I_n \mid O_{nx1}
\end{bmatrix}, \quad j = 1, 2, \ldots, p+1,
\]

\[
M_j = \begin{bmatrix}
    b_j & I_n \mid O_{nx1} \\
    a_j & I_n \mid O_{nx1}
\end{bmatrix}, \quad j = p+2, \ldots, n+1.
\]

By induction we easily verify

\[
\begin{bmatrix}
    b_{p+1} \\
    I_n \mid O_{nx1}
\end{bmatrix} \ldots \begin{bmatrix}
    b_1 \\
    I_n \mid O_{nx1}
\end{bmatrix} = \begin{bmatrix}
    B_{p+1} \\
    I_{n-p} \mid O_{(n-p)\times(p+1)}
\end{bmatrix},
\]

where \( B_{p+1} = \begin{bmatrix} 2^p & 0 & \ldots & 0 \\ 2^{p-1} & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \ldots & 0 & 0 \end{bmatrix} \) is a matrix of dimension \((p + 1) \times (n + 1)\). In view of (17) we have

\[
\begin{bmatrix}
    a_{n+1} \\
    I_n \mid O_{nx1}
\end{bmatrix} \ldots \begin{bmatrix}
    a_{p+2} \\
    I_n \mid O_{nx1}
\end{bmatrix} = \begin{bmatrix}
    A_{n-p}(p + 2) \\
    I_{p+1} \mid O_{(p+1)\times(n-p)}
\end{bmatrix},
\]

where \( A_{n+1}(j) \) is defined in (18). Therefore

\[
M = M_{n+1} \cdots M_1 = \begin{bmatrix} 2^n & 0 & \ldots & 0 \\ 2^{n-1} & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 1 & \ldots & 0 & 0 \end{bmatrix}, \quad r = \rho(M) = 2^n. \square
\]

We conclude with a remark that similar families of methods can be defined for Newton’s type methods using Hermite’s interpolating polynomials. HMM can again be used for the convergence analysis of these methods.

4.2. Families of methods with memory

In this section we will use HMM to compare order of convergence of multipoint methods with memory. We are exploring root-finders for a nonlinear equation \( f(x) = 0 \) used to approximate an isolated simple real zero \( a \). We restrict our attention to methods that use Hermitian information: if the method uses \( f^{(j)}(x) \) at some point \( x \), function values \( f(x), f'(x), \ldots, f^{(r-1)}(x) \) are used as well. These type of methods were in detail discussed in [17] by Woźniakowski. He proved that for this type of methods those relayed on Hermite’s interpolation attain the highest order of convergence. Kung and Traub in [2] have proved that two types of sampling in interpolatory iteration functions (HIF for short) provide optimal order of convergence. These are

1° Newton’s type: \( f(y_1), f'(y_1), f(y_2), \ldots, f(y_n) \);

2° Steffensen’s type: \( f(y_0), f(y_1), f(y_2), \ldots, f(y_n) \),

with iteration index \( k \) omitted. For these reasons we explore convergence acceleration of optimal Newton’s type and Steffensen’s type methods and compare outcomes. Using HMM we will show that convergence
acceleration in HIF using information from one previous iteration gives better results for Steffensen’s type of methods.

Let \( \varphi_N \) and \( \varphi_S \) denote two optimal multipoint iterations of order \( 2^n \). \( \varphi_N \) will denote Newton’s type of methods, using sampling \( 1^n \), and \( \varphi_S \) Steffensen’s type of methods, using sampling \( 2^n \).

\[
\varphi_N : \begin{cases} 
y_{k,1} = x_k, \\
y_{k,j} = N_j(y_{k,j-1}, \ldots, y_{k,1}), & j = 2, \ldots, n+1, \\
x_{k+1} = y_{k,n+1}, & k \in \mathbb{N}_0. 
\end{cases}
\]

\[
\varphi_S : \begin{cases} 
y_{k,0} = x_k, \\
y_{k,1} = S_1(y_{k,0}) = y_{k,0} + \gamma f(y_{k,0}), \\
y_{k,j} = S_j(y_{k,j-1}, \ldots, y_{k,1}, y_{k,0}), & j = 2, \ldots, n+1, \\
x_{k+1} = y_{k,n+1}, & k \in \mathbb{N}_0. 
\end{cases}
\]

Therefore, \( \varphi_N = N_{n+1} \circ \cdots \circ N_2 \), and \( \varphi_S = S_{n+1} \circ \cdots \circ S_1 \) in the context of composition of iteration functions.

We wish to calculate the highest order of convergence obtainable for methods \( \varphi_N \) and \( \varphi_S \) when they are accelerated using information from one previous iteration. Modifications with memory of methods \( \varphi_N \) and \( \varphi_S \) will be denoted by \( \tilde{\varphi}_N \) and \( \tilde{\varphi}_S \), respectively. We are examining the case where all information from the current and one previous iteration are used in each step. Let the modifications with memory read

\[
\tilde{\varphi}_N : \begin{cases} 
y_{k,1} = x_k, \\
y_{k,j} = \tilde{N}_j(y_{k,j-1}, \ldots, y_{k,1}), & j = 2, \ldots, n+1, \\
x_{k+1} = y_{k,n+1}, & k \in \mathbb{N}_0, 
\end{cases}
\]

\[
\tilde{\varphi}_S : \begin{cases} 
y_{k,0} = x_k, \\
y_{k,1} = \tilde{S}_1(y_{k,0}) = y_{k,0} + \gamma f(y_{k,0}), \\
y_{k,j} = \tilde{S}_j(y_{k,j-1}, \ldots, y_{k,1}, y_{k,0}), & j = 1, \ldots, n+1, \\
x_{k+1} = y_{k,n+1}, & k \in \mathbb{N}_0. 
\end{cases}
\]

It is assumed that corresponding error relations of each step in \( \tilde{\varphi}_N \) and \( \tilde{\varphi}_S \) are of HIF form

\[
\tilde{\varphi}_N : y_{k,j} - \alpha = \tilde{N}_j(y_{k,j-1}, \ldots, y_{k,1}) - \alpha
\]

\[
\sim (y_{k,1} - \alpha) \prod_{i=1}^{j-1} (y_{k,i} - \alpha) \prod_{i=1}^{n} (y_{k,j-1} - \alpha),
\]

\( j = 2, \ldots, n+1; \)

\[
\tilde{\varphi}_S : y_{k,j} - \alpha = \tilde{S}_j(y_{k,j-1}, \ldots, y_{k,0}) - \alpha
\]

\[
\sim \prod_{i=0}^{j} (y_{k,i} - \alpha) \prod_{i=0}^{n} (y_{k,j-1} - \alpha),
\]

\( j = 1, \ldots, n+1. \)

**Theorem 4.3.** \( r(\tilde{\varphi}_S) > r(\tilde{\varphi}_N). \)

**Proof.** Associated matrices for Newton’s type methods with memory \( \varphi_N \) are

\[
M(\tilde{N}_j) = \begin{bmatrix} c_j \\ I_{2n-1} \end{bmatrix},
\]

\[
\begin{cases} 
\text{for } j = 2, \ldots, n+1. 
\end{cases}
\]

where \( c_j = [1 \ldots 1 2 1 \ldots 1 2 0 \ldots 0] \), for \( j = 2, \ldots, n+1. \)
It is easily proved that for $1 \leq j \leq n-1$ the product of associated matrices reads

$$M(\widetilde{N}_2) \cdots M(\widetilde{N}_2) = \begin{bmatrix} C_j \\ I_{2n-j} \big| O_{(2n-j-1) \times 1} \end{bmatrix},$$

where $C_j = \begin{bmatrix} 2/c_2 \\ 2^{j-1}c_2 \\ \vdots \\ 2^0c_2 \end{bmatrix}$.

$C_j$ is the matrix of dimension $(j + 1) \times 2n$. Then

$$M(\widetilde{q}_N) = \begin{bmatrix} C_{n-1} \\ I_n \big| O_{n \times n} \end{bmatrix}.$$  

Since $(n + 1)$-dimension leading principal submatrix of $M(\widetilde{q}_N)$ is

$$\begin{bmatrix} 2^n & 2^{n-1} & \cdots & 2^{n-j} & \cdots & 2^0 \\ 2^{n-1} & 2^{n-2} & \cdots & 2^{n-j} & \cdots & 2^0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 2^0 & 2^0 & \cdots & 2^0 & \cdots & 2^0 \\ 1 & 0 & \cdots & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 2^{n-1} & 0 & \cdots & 0 \\ 0 & 2^{n-2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 2^0 \\ 0 & 1 & \cdots & 0 \end{bmatrix}$$

introduce $A = \begin{bmatrix} 2^{n-1} & 2^{n-2} & \cdots & 2^0 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix}^T$ and $B = \begin{bmatrix} 2 & 1 & \cdots & 1 & 2 \\ 1 & 0 & \cdots & 0 & 0 \end{bmatrix}$. The characteristic polynomial of $M(\widetilde{q}_N)$ is obtained using Laplace’s expansion and Theorem 2.9:

$$P_{\widetilde{q}_N}(\lambda) = \det(\lambda I_{2n} - M(\widetilde{q}_N)) = \lambda^{n-1} \det(\lambda I_{n+1} - AB) = \lambda^{2n-2} \det(\lambda I_2 - BA) = \lambda^{2n-2} \begin{vmatrix} \lambda - 3 \cdot 2^{n-1} + 1 & -2 \\ -2^{n-1} & \lambda \end{vmatrix}.$$  

This gives

$$r(\widetilde{q}_N) = \rho(M(\widetilde{q}_N)) = \frac{1}{2} \left( 3 \cdot 2^{n-1} - 1 + \sqrt{1 + 2^n + 9 \cdot 2^{2(n-1)}} \right).$$

With notation similar to previous section, with $a_j = \begin{bmatrix} 1 & \ldots & 1 & 0 & \ldots & 0 \end{bmatrix}$, associated matrices for Steffensen’s type of methods read

$$M(\widetilde{S}_j) = \begin{bmatrix} a_{n+j+1} \\ I_{2n+1} \big| O_{(2n+1) \times 1} \end{bmatrix} \in \mathcal{M}_{(2n+2)(2n+2)}, \ 1 \leq j \leq n+1.$$  

Proceeding in the same manner as earlier, we obtain

$$M(\widetilde{q}_S) = \begin{bmatrix} A_{n+1}(n+2) \\ I_{n+1} \big| O_{(n+1) \times (n+1)} \end{bmatrix},$$
where $A_{n+1}(n+2)$ is the $(n+1) \times (2n+2)$ matrix with $i$-th row equal to $2^{n+1-i}a_{r,2}$, that is

$$A_{n+1}(n+2) = \begin{bmatrix} 2^n a_{n+2} \\ 2^{n-1} a_{n+2} \\ \vdots \\ 2^0 a_{n+2} \end{bmatrix} = \begin{bmatrix} 2^n & 2^n & \cdots & 2^n \\ 2^{n-1} & 2^{n-1} & \cdots & 2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2^{n-1} \\ \vdots \\ 2^0 \end{bmatrix}.$$

Since

$$\begin{bmatrix} 2^n & 2^n & \cdots & 2^n \\ 2^{n-1} & 2^{n-1} & \cdots & 2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \end{bmatrix},$$

we get $P_{\widehat{\psi}}(\lambda) = A^{2n} \begin{bmatrix} \lambda - 2^n + 1 \\ -2^n \end{bmatrix}^{-1}$. Therefore

$$r(\widehat{\psi}_S) = \rho(M(\widehat{\psi}_S)) = \frac{1}{2} \left( 2^{n+1} - 1 + \sqrt{2^{2(n+1)} + 1} \right).$$

Having in mind

$$r(\widehat{\psi}_N) = \frac{1}{2} \left( 3 \cdot 2^{n-1} - 1 + \sqrt{(3 \cdot 2^{n-1} - 1)^2 + 2^{n+2}} \right),$$

$$r(\widehat{\psi}_S) = \frac{1}{2} \left( 4 \cdot 2^{n-1} - 1 + \sqrt{(4 \cdot 2^{n-1} - 1)^2 + 2^{n+2}} \right),$$

we conclude that Steffensen’s type methods produce better order of convergence with the same information volume. □

Note that similar analysis can be performed for different type of memory usage and different error relations.

References