ON THE CONVERGENCE ORDER OF
THE MULTISTEP METHODS

Ion Păvăloiu

1 Introduction

In this paper we analyse some aspects concerning the convergence order of the
iterative methods of interpolatory type for solving scalar equations. A unitary
approach to these methods will enable us to analyse them and then to select
among them those with the optimal convergence order.

2 Convergence order

Denote $I = [a, b]$, $a, b \in \mathbf{R}$, $a < b$ and consider the equation

$$(2.1) \qquad\qquad f(x) = 0$$

where $f : I \to \mathbf{R}$. For the sake of the simplicity we shall suppose in the following
that equation (2.1) has a unique solution $\bar{x} \in I$. Let $g : I \to I$ be a function for
which $\bar{x}$ is the unique fixed point on I.

For the approximation of the solution of (2.1) there is generally considered,
under certain conditions, a sequence $(x_p)_{p \geq 0}$ generated by:

$$(2.2) \qquad\qquad x_{s+1} = g(x_s), \quad s = 0, 1, \ldots, \quad x_0 \in I$$

More generally, if $G : I^k \to I$ is a function of k variables for which the
restriction to the diagonal of the set I^k coincides with g, namely

$$G(x, x, \ldots, x) = g(x), \quad \text{for all } x \in I,$$

then there is considered the following iterative method:

$$(2.3) \quad x_{s+k} = G(x_s, x_{s+1}, \ldots, x_{s+k-1}), \quad s = 0, 1, \ldots, \quad x_0, x_1, \ldots, x_{k-1} \in I.$$

The convergence of the sequence $(x_s)_{s\geq 0}$ generated by (2.2) or (2.3) depends on certain properties of the functions f and g, respectively G. The amount of time needed to obtain a convergent approximation z_s of $\overline{z}$ depends both on the convergence order of the sequence $(x_p)_{p\geq 0}$ and on the amount of elementary operations that must be performed by the computer at each iteration step. In this paper we shall study only the first aspect of this problem, i.e. we shall deal with the convergence order.

We shall adopt as the convergence order a definition a little more generally then the one given by Ostrowski [3].

Consider an arbitrary sequence $(x_p)_{p\geq 0}$ satisfying, together with f and g, the following conditions:

a) $x_s \in I$ and $g(x_s) \in I$ for $s = 0, 1, \dots$;

b) the sequences $(x_p)_{p\geq 0}$ and $(g(x_p))_{p\geq 0}$ converge to $\overline{x}$, the solution of equation (2.1);

c) there exists $m \in \mathbb{R}$, $m > 0$ such that $0 < \|[x, y; f]\| \leq m$, for all $x, y \in I$, where we have denoted by $[x, y; f]$ the first order divided difference of f on the points x and y.

d) f is differentiable at $\overline{x}$.

Definition 2.1 The sequence $(x_p)_{p\geq 0}$ has the convergence order $\omega \in \mathbb{R}$, $\omega \geq 1$ with respect to the function g, if there exists the limit:

$$(2.4) \qquad \alpha = \lim_{p \to \infty} \frac{\ln |g(x_p) - \overline{x}|}{\ln |x_s - \overline{x}|}$$

and $\alpha = \omega$.

Remark 2.1 If the elements of $(x_p)_{p\geq 0}$ are generated by the iterative method (2.2), then the convergence order defined above coincides with the one defined in [3].

For the determination of the convergence order of some classes of iterative methods we shall need in the following two lemmas:

Lemma 2.1 If the sequence $(x_p)_{p\geq 0}$ and the functions f and g satisfy the properties a) - d) then the necessary and sufficient condition for the sequence $(x_p)_{p\geq 0}$ to have the convergence order $\omega \in \mathbb{R}$, $\omega \geq 1$ with respect to the function g is that the following limit to exist:

$$(2.5) \qquad \lim_{p \to \infty} \frac{\ln |f(g(x_p))|}{\ln |f(x_p)|} = \beta,$$

and $\beta = \omega$.

Proof. Supposing that one of the relations (2.4) or (2.5) holds and, taking into account the properties a) – d), we obtain:

$$\alpha = \lim \frac{\ln |g(x_\nu) - \overline{x}|}{\ln |z_\nu - \overline{x}|} = \lim \frac{\ln |f(g(x_\nu))| - \ln |[g(x_\nu), \overline{x}; f]|}{\ln |f(x_\nu)| - \ln |[x_\nu, \overline{x}; f]|} =$$

$$= \lim \frac{\ln |f(g(x_\nu))|}{\ln |f(x_\nu)|} \cdot \frac{1 - \dfrac{\ln |[g(x_\nu), \overline{x}; f]|}{\ln |f(g(x_\nu))|}}{1 - \dfrac{\ln |[x_\nu, \overline{x}; f]|}{\ln |f(x_\nu)|}} = \lim \frac{\ln |f(g(x_\nu))|}{\ln |f(x_\nu)|} = \overline{\alpha}. \quad \square$$

Lemma 2.2 *If $(u_\nu)_{\nu \geq 0}$ is a sequence of real numbers that satisfies:*

i. the sequence $(u_\nu)_{\nu \geq 0}$ is convergent and $\lim u_\nu = 0$;

ii. there exist the nonnegative numbers $\alpha_1, \alpha_2, \dots, \alpha_{n+1}$ and a bounded sequence $(c_\nu)_{\nu \geq 0}$, $c_\nu > 0$, $\nu = 0, 1, \dots$ for which $\inf \{c_\nu\} > 0$ such that

$$(2.6) \qquad u_{s+n+1} = c_s u_s^{\alpha_1} u_{s+1}^{\alpha_2} \dots u_{s+n}^{\alpha_{n+1}}, \quad s = 0, 1, \dots;$$

iii. the sequence $\left(\dfrac{\ln u_{\nu+1}}{\ln u_\nu} \right)_{\nu \geq 0}$ is convergent and $\lim \dfrac{\ln u_{\nu+1}}{\ln u_\nu} = \omega$.

Then ω is the positive root of the equation:

$$t^{n+1} - \alpha_{n+1} t^n - \alpha_n t^{n-1} - \dots - \alpha_2 t - \alpha_1 = 0.$$

Proof. From (2.6) we obtain

$$\lim_{s \to \infty} \frac{\ln u_{n+s+1}}{\ln u_{s+s}} = \lim_{s \to \infty} \frac{\ln c_s}{\ln u_{n+s}} + \sum_{i=0}^{n} \alpha_{i+1} \lim_{s \to \infty} \frac{\ln u_{s+i}}{\ln u_{s+i}}.$$

Taking into account the equalities

$$\lim_{s \to \infty} \frac{\ln c_s}{\ln u_{n+s}} = 0 \text{ and } \lim_{s \to \infty} \frac{\ln u_{s+i}}{\ln u_{s+i}} = \frac{1}{\omega^{n-i}}, \quad i = \overline{0, n}$$

we obviously have

$$\omega = \sum_{i=0}^{n} \alpha_{i+1} \frac{1}{\omega^{n-i}},$$

i.e.,

$$\omega^{n+1} - \alpha_{n+1} \omega^n - \alpha_n \omega^{n-1} - \dots - \alpha_2 \omega - \alpha_1 = 0. \quad \square$$

In the following we shall consider the equations:

$$(2.7) \qquad P(t) = t^{n+1} - \alpha_{n+1} t^n - \alpha_n t^{n-1} - \dots - \alpha_2 t - \alpha_1 = 0;$$

$$(2.8) \qquad Q(t) = t^{n+1} - \alpha_1 t^n - \alpha_2 t^{n-1} - \ldots - \alpha_n t - \alpha_{n+1} = 0;$$

$$(2.9) \qquad R(t) = t^{n+1} - \alpha_{i_1} t^n - \alpha_{i_2} t^{n-1} - \ldots - \alpha_{i_n} t - \alpha_{i_{n+1}} = 0,$$

where we shall suppose that $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ satisfy

$$(2.10) \qquad \alpha_i \geq 0 \quad i = \overline{1, n+1}, \quad \sum_{i=1}^{n+1} \alpha_i > 1$$

$$(2.11) \qquad \alpha_{n+1} \geq \alpha_n \geq \ldots \geq \alpha_2 \geq \alpha_1.$$

$i_1, i_2, \ldots, i_{n+1}$ being an arbitrary permutation of the numbers $1, 2, \ldots, n+1$.

Lemma 2.3 *If $\alpha_1, \alpha_2, \ldots, \alpha_{n+1}$ satisfy conditions (2.10) then any of the equations (2.9) has a unique solution $\omega > 1$. Moreover, if (2.11) holds and we denote by a, b, c the positive solutions of (2.7), (2.8) and (2.9) then*

$$(2.12) \qquad 1 < b \leq c < a.$$

i.e. equation (2.7) has the greatest positive solution.

Proof.   Consider one of the $(n+1)!$ equations of the form (2.9), denote by s the greatest natural number for which $\alpha_{i_s} \neq 0$, i.e. $\alpha_{i_{s+1}} = \alpha_{i_{s+2}} = \ldots = \alpha_{i_{n+1}} = 0$ and consider the function $\varphi(t) = \dfrac{R(t)}{t^{n-s+1}}$.

We have $\varphi(1) = 1 - \alpha_{i_1} - \alpha_{i_2} - \ldots - \alpha_{i_s} < 0$ and $\lim_{t \to \infty} \varphi(t) = +\infty$, whence it follows that the equation $R(t) = 0$ has a solution greater than 1. Since the function $\varphi(t) = \dfrac{R(t)}{t^{n+1}}$ is increasing in the variable $x = \frac{1}{t}, x > 0$ it follows that this solution is unique.

In order to prove (2.12) it will suffice to show that $R(b) \leq 0$ and $R(a) \geq 0$. Indeed, since $Q(b) = 0$, we have:

$$R(b) = R(b) - Q(b) = (\alpha_2 - \alpha_{i_1})b^n + (\alpha_3 - \alpha_{i_2})b^{n-1} + \ldots + (\alpha_n - \alpha_{i_s})b + \alpha_{n+1} - \alpha_{i_{n+1}} =$$
$$= -(b-1)[(\alpha_1 - \alpha_{i_1})b^{n-1} + (\alpha_1 + \alpha_2 - \alpha_{i_1} - \alpha_{i_2})b^{n-2} + \ldots$$
$$+ (\alpha_1 + \alpha_2 + \ldots + \alpha_{n-1} - \alpha_{i_1} - \alpha_{i_2} - \ldots - \alpha_{i_{n-1}})b +$$
$$+ (\alpha_1 + \alpha_2 + \ldots + \alpha_n - \alpha_{i_1} - \alpha_{i_2} - \ldots - \alpha_{i_n})],$$

whence it follows that $R(b) < 0$, since $b > 1$ and the inequalities

$$\alpha_1 + \alpha_2 + \ldots + \alpha_s - \alpha_{i_1} - \alpha_{i_2} - \ldots - \alpha_{i_s} \leq 0, \quad s = \overline{1, n}$$

hold.

The inequality $R(a) \geq 0$ is proved in a similar manner.□

3   Iterative methods of interpolatory type

It is known that most of the iterative methods are obtained by inverse interpolation of Lagrange, Taylor or Hermite type.

Let $F = f(I)$ be the image of I by f. We shall suppose that f is derivable on I up to the order $n + 1$, and $f'(x) \neq 0$ for all $x \in I$. It then follows that $f : I \to F$ is invertible, so there exists $f^{-1} : F \to I$. It is obvious that if $\overline{x}$ is the solution of (2.1) then

(3.1)
$$x = f^{-1}(0).$$

In order to determine an approximation $\dot{x}$ of $\overline{x}$ it is sufficient to determine a function h which approximates f^{-1} on a neighborhood V_0 of $y = 0$.

If

(3.2)
$$f^{-1}(y) = h(y) + R\left(f^{-1}, y\right), \quad \forall y \in V_0,$$

then we can consider that $\dot{x} \approx \overline{x}$ with the error

(3.3)
$$|x - \dot{x}| = \left| R\left(f^{-1}, 0\right)\right|, \quad \text{where } \dot{x} = h(0).$$

A simple and efficient method for constructing functions that approximate f^{-1} on V_0 is the inverse interpolation. The most general method of this type is the method that lead us to the Hermite inverse interpolatory polynomial.

In order to construct the Hermite inverse interpolatory polynomial in its most general form we must know both the values of f^{-1} and the values of the derivatives of f^{-1} at some precized points from V_0.

Concerning the successive derivatives of f^{-1} we shall rely on the following lemma.

Lemma 3.1 [6]   *If $f : I \to F$ is derivable up to the order $n + 1$ and $f'(x) \neq 0$ for all $x \in I$, then there exists $f^{-1} : F \to I$ and the following equality holds:*

(3.4)
$$\left(f^{-1}\right)^{(k)}(y) = \sum \frac{(2k - i_1 - 2)! \, (-1)^{k+i_1-1}}{i_2! i_3! \ldots i_k! \, [f'(x)]^{2k-1}} \prod_{j=1}^{k} \left(\frac{f^{(j)}(x)}{j!}\right)^{i_j}.$$

where $y = f(x)$ and the above sum extends on all nonnegative integer solutions of the system

$$\begin{cases} i_2 + 2i_3 + 3i_4 + \ldots + (k-1)i_k = k - 1 \\ i_1 + i_2 + \ldots + i_k = k - 1. \end{cases}$$

Let $z_i \in I$, $i = \overline{1, n+1}$ be $n + 1$ interpolation nodes and $a_1, a_2, \ldots, a_{n+1} \geq 1$, $n + 1$ natural numbers. Denote by $m + 1$ their sum:

(3.5)
$$a_1 + a_2 + \ldots + a_{n+1} = m + 1.$$

We also denote $y_i = f(x_i)$, $i = \overline{1, n+1}$ and so $f^{-1}(y_i) = x_i$.

For the construction of the Hermite inverse interpolatory polynomial we consider as interpolation nodes the numbers $y_i \in F$, $i = \overline{1, n+1}$, and we need a polynomial

$$H\left(y_1, \alpha_1; y_2, \alpha_2; \ldots; y_{n+1}, \alpha_{n+1}; f^{-1} \mid y\right) = H(y)$$

which satisfies

(3.6) $$H^{(j)}(y_i) = \left(f^{-1}\right)^{(j)}(y_i), \quad j = \overline{0, \alpha_i - 1}, \quad i = \overline{1, n+1}.$$

It is well known that such a polynomial has the following form

$$H\left(y_1, \alpha_1; y_2, \alpha_2; \ldots; y_{n+1}, \alpha_{n+1}; f^{-1} \mid y\right) =$$

(3.7)

$$= \sum_{i=1}^{n+1} \sum_{j=0}^{\alpha_i - 1} \sum_{k=0}^{j-1} \left(f^{-1}\right)^{(j)}(y_i) \frac{1}{k!j!} \left[\frac{(y-y_i)^{\alpha_i}}{\omega(y)}\right]_{y=y_i}^{(k)} \cdot \frac{\omega(y)}{(y-y_i)^{\alpha_i - k - j}},$$

where

(3.8) $$\omega(y) = \prod_{i=1}^{n+1} (y - y_i)^{\alpha_i}.$$

The following equality holds:

(3.9) $$f^{-1}(y) = H\left(y_1, \alpha_1; y_2, \alpha_2; \ldots; y_{n+1}, \alpha_{n+1}; f^{-1} \mid y\right) + R\left(f^{-1}, y\right)$$

where

(3.10) $$R\left(f^{-1}, y\right) = \frac{\left(f^{-1}\right)^{(m+1)}(\theta)}{(m+1)!} \omega(y),$$

θ being a point belonging to the smallest interval determined by the points $y, y_1, y_2, \ldots, y_{n+1}$.

From (3.7), for $\alpha_1 = \alpha_2 = \ldots = \alpha_{n+1} = 1$, there is obtained the Lagrange inverse interpolatory polynomial, i.e.

$$L\left(y_1, y_2, \ldots, y_{n+1}; f^{-1}y\right) = \sum_{i=1}^{n+1} \frac{\omega(y)}{(y - y_i)\omega'(y_i)},$$

and for $n = 0$, again (3.7) gives the Taylor inverse interpolatory polynomial:

$$T\left(y_1, f^{-1}y\right) = \sum_{j=0}^{m} \frac{1}{j!} \left(f^{-1}\right)^{(j)}(y_1)(y - y_1)^j.$$

Let $x_s, x_{s+1}, ..., x_{s+n} \in I$ be $n+1$ approximations for the solution $\bar{z}$ of equation (2.1). Then another approximation x_{s+n+1} can be obtained by

$$(3.11) \quad x_{s+n+1} = H\left(y_s, a_1; y_{s+1}, a_2; ...; y_{s+n}, a_{n+1}; f^{-1} \mid 0\right), \quad s = 1, 2, ...$$

where the function H is given by (3.7) and the polynomial ω_s is given by

$$\omega_s(y) = \prod_{i=s}^{n+s} (y - y_i)^{m_i}, \quad s = 1, 2, ...$$

It can be easily seen that, taking into account (3.9),

$$(3.12) \quad |f(x_{n+s+1})| = |f'(\beta_s)| \frac{\left|(f^{-1})^{n+1}(\theta_s)\right|}{(m+1)!} \prod_{i=0}^{n} |f(x_{s+i})|^{n+1}, \quad s = 1, 2, ...,$$

if the equence $\{x_s\}_{p>0}$ is generated by (3.11). For all $s \geq 1$, the numbers θ_s belong to the smallest interval determined by $0, y_s, y_{s+1}, ..., y_{s+n}$ and β_s are numbers belonging to the open intervals determined by $\bar{z}$ and x_{s+n+1}.

Supposing that $c_s = |f'(\beta_s)| \frac{(f^{-1})^{(m+1)}(\theta_s)}{(m+1)!}$, $s = 1, 2, ...$, verify the hypotheses of Lemma 2.2 and, moreover, $\lim_{s \to \infty} f(x_s) = 0$, then, taking into account Lemma 2.1, it follows that the convergence order of method (3.11) is given by the positive solution of equation

$$(3.13) \qquad t^{s+1} - a_{n+1} t^n - a_n t^{n-1} - ... - a_2 t - a_1 = 0.$$

If $a_1 = a_2 = ... = a_{n+1} = q$, then the corresponding iterative method from (3.11) has the convergence order $\omega_{n+1}(q)$ given by the positive solution of the equation:

$$t^{n+1} - q t^n - q t^{n-1} - ... - qt - q = 0.$$

In [5] there is shown that the number $\omega_n(q)$ satisfies

a') $\omega_n(q) < \omega_{n+1}(q)$, $n = 1, 2, ...$;

b') $\max\left\{q, \frac{q+1}{q+2}(q+1)\right\} < \omega_{n+1}(q) < q + 1$;

c') $\lim \omega_n(q) = q + 1$.

For $q = 1$, the properties a') – c') can be found in [3]

For $n = 0$, the corresponding iterative method has the form

$$x_{s+1} = x_s - \frac{(f^{-1})'(y_s)}{1!} f(x_s) + \frac{1}{2!} (f^{-1})''(y_s) f^2(x_s) + \ldots$$

$$+ (-1)^m \frac{1}{m!} (f^{-1})^{(m)}(y_s) f^m(x_s), \quad s = 1, 2, \ldots,$$

i.e., the Chebyshev method. It can be easily seen that the convergence order of this method is $m + 1$.

We propose ourselves to find bounds for the convergence order ω of (3.11), i.e. for the solutions of (3.13). First we shall prove the following lemma.

Lemma 3.2 *The positive root ω of equation (3.11) satisfies*

(3.14)
$$(m+1)^n \leq \omega \leq 1 + \max_{1 \leq i \leq n} \{a_i\},$$

where

$$a = \frac{m+1}{(n+1)(m+1) - \sum_{i=1}^{n+1} (i-1) a_i}$$

Proof. Let $\beta = (m+1)^n$ and $P_{n+1}(t) = t^{n+1} - a_{s+1} t^n - \ldots - a_2 t - a_1$. For the first part of inequality (3.14) it will suffice to show that $P_{n+1}(\beta) \leq 0$. Using the inequality between the geometric and the arithmetic mean, i.e.

$$\frac{\sum_{i=1}^{n+1} p_i a_i}{\sum_{i=1}^{n+1} p_i} \geq \left(\prod_{i=1}^{n+1} a_i^{p_i} \right)^{\frac{1}{\sum_{i=1}^{n+1} p_i}}, \quad a_i > 0, \quad p_i \geq 0, \quad i = \overline{1, n+1}, \quad \sum_{i=1}^{n+1} p_i > 0,$$

we obtain

$$P_{n+1}(\beta) = \beta^{n+1} - \sum_{i=1}^{n+1} a_i \beta^{i-1} = \beta^{n+1} - \frac{\sum_{i=1}^{n+1} a_i \beta^{i-1}}{\sum_{i=1}^{n+1} a_i} \cdot \sum_{i=1}^{n+1} a_i \leq$$

$$\leq \beta^{n+1} - \left(\sum_{i=1}^{n+1} a_i \right) \left(\prod_{i=1}^{n+1} \beta^{(i-1)a_i} \right)^{\frac{1}{\sum_{i=1}^{n+1} a_i}} = \beta^{n+1} - (m+1) \left(\prod_{i=1}^{n+1} \beta^{(i-1)a_i} \right)^{\frac{1}{n+1}} = 0.$$

for $\beta = (m+1)^n$.

For the second inequality in (3.14) it easily comes that $P_{n+1}(a) \geq 0$ for $a = 1 + \max_{1 \leq i \leq n+1} \{a_i\}$.

4 Interpolatory methods having optimal convergence order

To each permutation of the set $\{1, 2, ..., n+1\}$ there corresponds an iterative method of the form

$$(4.1) \quad \begin{aligned} x_{n+2} &= H\left(y_{a_1}, a_{i_1}; y_{i_2}, a_{i_2}; ... y_{i_{n+1}}, a_{i_{n+1}}; f^{-1}0\right) \\ x_{n+s+2} &= H\left(y_{i_1+s}, a_{i_1}; y_{i_2+s}, a_{i_2}; ... y_{i_{n+1}+s}, a_{i_{n+1}}; f^{-1}0\right), \quad s = 1, 2, ... \end{aligned}$$

There are $(n+1)!$ iterative methods. Lemma 2.3 offers the possibility to establish which method of the form (4.1) has the optimal convergence order.

Theorem 4.1 *Among the $(n+1)!$ iterative methods of the form (4.1) the method with the highest convergence order is the one for which*

$$a_{i_1} \le a_{i_2} \le ... \le a_{i_{n+1}}.$$

The proof of the above theorem is obtained as a directly consequence of Lemma 2.3.

In the following we shall apply the results of Theorem 4.1 to a particular case. For $n = 1$, by (3.11) we obtain two iterative methods:

$$(4.2) \quad \begin{aligned} x_3 &= H\left(y_1, a_1; y_2, a_2; f^{-1}0\right), x_1, x_2 \in I, y_1 = f(x_1), y_2 = f(x_2) \\ x_{n+1} &= H\left(y_{n-1}, a_1; y_n, a_2; f^{-1}0\right), \quad n = 3, 4, ... \end{aligned}$$

and

$$(4.3) \quad \begin{aligned} x_3 &= H\left(y_1, a_2; y_2, a_1; f^{-1}0\right), x_1, x_2 \in I, y_1 = f(x_1), y_2 = f(x_2) \\ x_{n+1} &= H\left(y_{n-1}, a_2; y_n, a_1; f^{-1}0\right), \quad n = 3, 4. \end{aligned}$$

The convergence order ω_1 of method (4.2) is given by the positive solution of

$$t^2 - a_2 t - a_1 = 0$$

and for (4.3) the convergence order is the solution of

$$t^2 - a_1 t - a_2 = 0.$$

It is easily seen that if $a_2 \ge a_1$ then $\omega_2 \le \omega_1$, and so the iterative method (4.2) has a greater convergence order than the one of (4.3).

References

[1] R. Brent, S. Winograd, F. Wolfe, *Optimal Iterative Processes for Root-Finding.*, Numer. Math. **20** (1973), 327-341.

[2] H.T. Kung, J.F. Traub, *Optimal Order and Efficiency for Iterations with Two Evaluations*, Numer. Anal., Vol. **13**, 1, (1976), 84-99.

[3] A.M. Ostrowski, *Solution of Equations and Systems of Equations*, Academic Press, New York and London, (1966).

[4] I. Păvăloiu, *Optimal Problems Concerning Interpolation Methods of Solution of Equations*, Publ. Inst. Math., **52** (66) (1992), 113-126.

[5] J.F. Traub , *Iterative Methods for Solution of Equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, (1964).

[6] B.A. Turowicz, *Sur les derivées d'ordre supérieur d'une fonction inverse*, Ann. Polon. Math. **8** (1960), 265-269.