REVUE D'ANALYSE NUMÉRIQUE ET DE THÉORIE DE L'APPROXIMATION
Rev. Anal. Numér. Théor. Approx., vol. 35 (2006) no. 1, pp. 87–94
ictp.acad.ro/jnaat

ON A STEFFENSEN-HERMITE TYPE METHOD FOR APPROXIMATING THE SOLUTIONS OF NONLINEAR EQUATIONS*

ION PĂVĂLOIU†

Abstract. It is well known that the Steffensen and Aitken-Steffensen type methods are obtained from the chord method, using controlled nodes. The chord method is an interpolatory method, with two distinct nodes. Using this remark, the Steffensen and Aitken-Steffensen methods have been generalized using interpolatory methods obtained from the inverse interpolation polynomial of Lagrange or Hermite type. In this paper we study the convergence and efficiency of some Steffensen type methods which are obtained from the inverse interpolatory polynomial of Hermite type with two controlled nodes.

MSC 2000. 65H05.
Keywords. Nonlinear equations in $\mathbb{R}$, Steffensen and Aitken-Steffensen methods, inverse interpolatory polynomial of Hermite type.

1. INTRODUCTION

The most well-known methods for approximating the solutions of nonlinear equations are of interpolatory type.

The Newton type methods and, more generally, the Chebyshev type methods are obtained from the inverse interpolatory polynomial of Taylor type, with one node. The chord method and its generalizations are obtained from the inverse interpolation polynomial of Lagrange, or more generally, from the inverse interpolation polynomial of Hermite [4], [7], [10]. The Steffensen and Aitken-Steffensen methods are also of interpolatory type, but their nodes are controlled at each step [6]–[9]. For the methods of interpolatory type it is necessary to compute the values of the derivative of the inverse function at different points in $\mathbb{R}$.

Let $f : [c, d] \longrightarrow \mathbb{R}$ be a given function, where $c, d \in \mathbb{R}$, $c < d$, and denote by $F = f([c, d])$ the set of the values of f on $[c, d]$. Assume that f is one-to-one and therefore there exists $f^{-1} : F \longrightarrow [c, d]$.

Consider the equation

$$f(x) = 0. \tag{1}$$

If equation (1) has a solution $\bar{x} \in [c, d]$ then obviously

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

*This work has been supported by the Romanian Academy under grant GAR 14/2005.
†"T. Popoviciu" Institute of Numerical Analysis, P.O. Box 68-1, Cluj-Napoca, Romania, e-mail: pavaloiu@ictp.acad.ro, web: www.ictp.acad.ro/~pavaloiu.

It is therefore natural to seek methods of approximation of the values of f^{-1} at $y=0$, in order to determine different approximations to $\bar{x}$.

Several methods of approximation for $f^{-1}(0)$ have been studied in papers such as [2]–[5], [7], [8], [10].

In the following we shall consider a method of Hermite type with two interpolation nodes. Such a method has been studied in [2], [3], [8], etc.

Let $p,q\in\mathbb{N}$, $p,q\geq 1$, and $x_1,\ x_2\in[c,d]$. Denote by

$$H(y)=H(y_1,p;y_2,q;f^{-1}\mid y) \tag{2}$$

the Hermite polynomial associated to the inverse function f^{-1} with the multiple nodes: y_1 of order p, and y_2 of order q, where $y_i=f(x_i)$, $i=1,2$. In order that polynomial (2) exists, it suffices that function f^{-1} to be differentiable up to the order $p+q$ in F.

In this sense the following theorem holds [4], [11].

Theorem 1. *If f satisfies the following conditions:*

i$_1$. *$f:[c,d]\longrightarrow F$ is one-to-one;*
ii$_1$. *f is differentiable up to the order $n\in\mathbb{N}$ at each point $x\in[c,d]$;*
iii$_1$. *$f'(x)\neq 0$ for all $x\in[c,d]$,*

then the inverse function f^{-1} is differentiable up to the order n at each point $y\in F$, and the following relations hold:

$$[f^{-1}(y)]^{(k)}=\sum\tfrac{(2k-2-i_1)(-1)^{k-1+i_1}}{i_2!i_3!\ldots i_k![f'(x)]^{2k-1}}\left(\tfrac{f'(x)}{1!}\right)^{i_1}\cdot\left(\tfrac{f''(x)}{2!}\right)^{i_2}\cdots\left(\tfrac{f^{(k)}(x)}{k!}\right)^{i_k}, \tag{3}$$

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

$$\begin{matrix} i_2+2i_3+\cdots+(k-1)i_k=k-1\\ i_1+i_2+i_3+\cdots+i_k=k-1. \end{matrix} \tag{4}$$

Taking into account the above relations, under the hypothesis that f admits derivatives up to the order $p+q$ on $[c,d]$, $f'(x)\neq 0,\forall x\in[c,d]$, and using the remainder of the Hermite interpolation, it follows:

$$f^{-1}(y)=H(y)+\tfrac{[(f^{-1}(\eta)]^{(p+q)}}{(p+q)!}(y-y_1)^p(y-y_2)^q \tag{5}$$

for some $\eta\in\operatorname{int}(F)$.

Setting $y=0$ in (5), then we obtain for $\bar{x}$ the following relation:

$$\bar{x}=f^{-1}(0)=H(0)+\tfrac{[f^{-1}(\xi)]^{(p+q)}}{(p+q)!}(-1)^{p+q}y_1^py_2^q. \tag{6}$$

Denote by x_3 the next approximation for $\bar{x}$, obtained from (6)

$$x_3=H(y_1,p;y_2,q;f^{-1}|0). \tag{7}$$

If we assume that the derivative of order $p+q$ of the inverse function is bounded, i.e., there exists $M\in\mathbb{R},M>0$, such that

$$\left|[f^{-1}(y)]^{(p+q)}\right|\leq M,\ \text{for all } y\in F, \tag{8}$$

then by (2), (6) and (7) we deduce:

$$|\bar{x}-x_3|\leq\tfrac{M}{(p+q)!}\left|f(x_1)\right|^p\left|f(x_2)\right|^q.$$

In general, if $x_n, x_{n+1} \in [c,d]$ are two approximations for $\bar{x}$, then

(9) $$x_{n+2} = H(y_n, p; y_{n+1}, q; f^{-1} \mid 0), \quad n = 1, 2, \ldots,$$

where $y_n = f(x_n)$ while $y_{n+1} = f(x_{n+1})$ is a new approximation for $\bar{x}$, and obviously

(10) $$|\bar{x} - x_{n+2}| \leq \tfrac{M}{(p+q)!} \, |f(x_n)|^p \, |f(x_{n+1}|^q \,, \quad n = 1, 2, \ldots.$$

Taking into account all the above relations one can generate the sequence of approximations $(x_n)_{n\geq 1}$, under the assumption that at each iteration step, the new approximation x_{n+2} obtained by (9) from x_n and x_{n+1}, lies in $[c,d]$.

It can be easily seen that the convergence order of method (9) is given by the positive root of equation [1]–[4], [8], [10]:

$$t^2 - qt - p = 0$$

i.e.,

(11) $$\omega = \tfrac{q+\sqrt{q^2+4p}}{2}.$$

If we assume now that the Fréchet derivative f' of f satisfies

(12) $$\sup_{x\in[c,d]} |f'(x)| \leq a, \;\; a \in \mathbb{R}, a > 0,$$

then (10) becomes

(13) $$|\bar{x} - x_{n+2}| \leq \tfrac{Ma^{p+q}}{(p+q)!} \, |\bar{x} - x_{n+1}|^q \, |\bar{x} - x_n|^p \,, \quad n = 1, 2, \ldots.$$

In the following we shall denote ρ_i, $i = 1, 2, \ldots$, the expressions

(14) $$\rho_i = a \left[\tfrac{Ma}{(p+q)!} \right]^{\frac{1}{p+q-1}} |\bar{x} - x_i| \,.$$

By (13) and (14) we obtain

$$\rho_{n+2} \leq \rho_{n+1}^q \rho_n^p, \quad n = 1, 2, \ldots.$$

If we assume now that ρ_1 obeys

(15) $$\rho_1 < 1$$

and, moreover, $\rho_2 \leq \rho_1^\omega$, with ω given by (11), we easily deduce that for all $n \in \mathbb{N}, n \geq 2$ we have

$$\rho_n \leq \rho_1^{\omega^{n-1}}, \; n = 2, 3, \ldots$$

and

$$\lim \rho_n = 0$$

which implies

$$\lim_{n\to\infty} x_n = \bar{x}.$$

Theorem 2. *If the following conditions hold:*

i$_2$. *function f obeys the assumptions of Theorem 1 for $n = p + q$;*
ii$_2$. *equation* (1) *has a root $\bar{x} \in]c,d[$;*
iii$_2$. *$[f^{-1}(y)]^{(p+q)}$ verifies* (8);
iv$_2$. *$f'\,(x)$ verifies* (12);
v$_2$. *ρ_1 given by* (14) *verifies* (15) *and $\rho_2 \leq \rho_1^\omega$, where ω is given by* (11);
vi$_2$. *the sequence $(x_n)_{n\geq 1}$ generated by* (9) *remains in $[c,d]$,*

then the following relations hold:

j$_2$. $|\bar{x} - x_n| \le a^{-1}[\frac{Ma}{(p+q)!}]^{-\frac{1}{p+q-1}} \rho_1^{\omega^{n-1}}, \quad n = 1, 2, \dots;$
jj$_2$. $\lim x_n = \bar{x}$.

Remark 1.1. The uniqueness of the root $\bar{x}$ of equation (1) is ensured by hypothesis iii$_1$ from Theorem 1. □

We consider in the following, besides (1), an equivalent equation, of the form

$$x - \varphi(x) = 0 \tag{16}$$

where

$$\varphi : [c, d] \longrightarrow [c, d], \quad \varphi(\bar{x}) = \bar{x}.$$

We shall show in the following section that if instead of method (9) we consider the iterative method

$$x_{n+1} = H(f(x_n), p; f(\varphi(x_n)), q; f^{-1} \mid 0), \quad n = 1, 2, \dots, x_1 \in [c, d], \tag{17}$$

then the r-convergence order of this resulted method, which we call of Steffensen-Hermite type, is at least $p+q$, i.e., substantially higher than ω, given by (11).

We shall also determine the values of p and q for which the methods is the class (17) are optimal with respect to the efficiency index [4].

In section 3 we shall study the convergence of the optimal methods determined in Section 2.

2. THE CONVERGENCE AND THE EFFICIENCY OF THE STEFFENSEN-HERMITE TYPE METHODS

We shall assume that the function φ from (16) is Lipschitz, i.e., there exists $b \in \mathbb{R}$, $b > 0$, such that

$$|\varphi(x) - \varphi(y)| \le b\,|x - y|\,, \quad x, y \in [c, d]. \tag{18}$$

For the approximation of the solution $\bar{x}$ of (1), consider the sequence $(x_m)_{m \ge 1}$, generated by (17).

Concerning the convergence of this method the following result holds.

Theorem 3. *If the functions f and φ obey the following conditions:*

i$_3$. *function f obeys conditions* i$_2$–iv$_2$ *from Theorem* 2;
ii$_3$. *the approximation $x_1 \in [c, d]$ verifies*

$$\delta_1 = a\left[\frac{Mab^q}{(p+q)!}\right]^{\frac{1}{p+q-1}} |\bar{x} - x_1| < 1, \tag{19}$$

where b is the Lipschitz constant from relation (18) *and* $0 < b < 1$;
iii$_3$. *the elements of the sequence $(x_n)_{n \ge 1}$ generated by* (17) *remain in* $[c, d]$,

then the following relations hold:

j$_3$. $|\bar{x} - x_n| \le a^{-1}[\frac{Mab^q}{(p+q)!}]^{-\frac{1}{p+q-1}} \delta_1^{(p+q)^{(n-1)}};$
jj$_3$. $\lim x_n = \bar{x}$.

Proof. The proof can be done along the same lines as in Theorem 2. □

From the inequality given in j$_3$ one can see that the r-convergence order of method (16) is $p + q$.

The following function evaluations are required by method (17) at each iteration step n in order to determine x_{n+1}, under the hypothesis that function φ from (16) is expressed with the aid of function f and its derivatives up to the order $r = \min\{p, q\}$:

$$f(x_n), f'(x_n), \ldots, f^{(p-1)}(x_n),$$
$$f(\varphi(x_n), f'(\varphi(x_n)), \ldots, f^{(q-1)}(\varphi(x_n)),$$

i.e., $p + q$ function evaluations.

The values of the successive derivatives of f^{-1} at the points $y_n = f(x_n)$ and $\bar{y}_n = f(\varphi(x_n))$ respectively, are determined by (3).

We shall admit that the number of function evaluations which must be performed from iteration step n to $n+1$ is proportional to $p+q$, with a factor $\alpha \in \mathbb{R}$, $\alpha > 0$. In this case, the efficiency index (see, e.g., [4], [10]) of method (17) is given by relation

$$I(p, q) = (p + q)^{\frac{1}{\alpha(p+q)}}.$$

An elementary study on the function $g : (0, +\infty) \longrightarrow (0, +\infty)$ given by

$$g(t) = t^{\frac{1}{\alpha t}},$$

show that this function attains its maximum at $\bar{t} = e$, and also that g is increasing on $(0, e)$ and decreasing on $(e, +\infty)$, i.e. $\bar{t}$ is a unique stationary point.

From here it follows that

$$I_3(p, q) = [\sqrt[3]{3}]^{\frac{1}{\alpha}}$$

while for $p + q = 2$

$$I_2(p, q) = [\sqrt[2]{2}]^{\frac{1}{\alpha}}.$$

It is clear that $I_3(p, q) > I_2(p, q)$ and therefore $I(p, q)$ has the maximum value for $p + q = 3$.

We shall point out and study the optimal methods in the following section.

3. THE CONVERGENCE OF THE OPTIMAL METHODS

We shall determine in the beginning the methods which correspond to the two cases which may be considered for $p + q = 3$, i.e., $p = 1$, $q = 2$ or $p = 2$, $q = 1$.

In the first case ($p = 1$, $q = 2$) we want to determine a polynomial H of the form:

$$H(y) = a_0 + a_1 y + a_2 y^2$$

with the conditions:

$$\begin{aligned} &H(y_1) = x_1, \quad x_1 \in [c, d], \\ &H(y_2) = \varphi(x_1), \\ &H'(y_2) = \tfrac{1}{f'(\varphi(x_1))} = [f^{-1}(y_2)]', \end{aligned} \tag{20}$$

where $y_1 = f(x_1)$ and $y_2 = f(\varphi(x_1))$.

Conditions (20) lead to a system of 3 equations with the unknowns a_0, a_1, a_2. For our purpose, of approximating the solution of (1), we are interested only in the value $H(0) = a_0$. Denoting by $[\varphi(x_1), \varphi(x_1); f] = f'(\varphi(x_1))$, then for a_0 we obtain the following two equivalent expressions:

$$\begin{aligned} a_0 &= \varphi(x_1) - \tfrac{f(\varphi(x_1))}{[\varphi(x_1),\varphi(x_2);f]} - \tfrac{[x_1,\varphi(x_1),\varphi(x_1);f]}{[\varphi(x_1),\varphi(x_1);f][x_1,\varphi(x_1);f]^2} f^2(\varphi(x_1)) \\ &= x_1 - \tfrac{f(x_1)}{[x_1,\varphi(x_1);f]} - \tfrac{[x_1,\varphi(x_1),\varphi(x_1);f]}{[\varphi(x_1),\varphi(x_1);f][x_1,\varphi(x_1);f]^2} f(x_1) f(\varphi(x_1)). \end{aligned}$$

The above expressions for a_0 lead us to the sequences $(x_n)_{n\geq 1}$ and $(\varphi(x_n))_{n\geq 1}$, generated by

$$x_{n+1}=x_n-\tfrac{f(x_n)}{[x_n,\varphi(x_n);f]}-\tfrac{[x_n,\varphi(x_n),\varphi(x_n);f]}{[\varphi(x_n),\varphi(x_n);f][x_n,\varphi(x_n);f]^2}f(x_n)f\left(\varphi(x_n)\right),\quad n=1,2,\ldots \tag{21}$$

or equivalently

$$x_{n+1}=\varphi(x_n)-\tfrac{f(\varphi(x_n))}{[\varphi(x_n),\varphi(x_n);f]}-\tfrac{[x_n,\varphi(x_n),\varphi(x_n);f]}{[\varphi(x_n),\varphi(x_n);f][x_n,\varphi(x_n);f]^2}f^2(\varphi(x_n)),\quad n=1,2,\ldots. \tag{22}$$

In the second case, $p=2$, $q=1$, we obtain in our analogous fashion, the following expressions for $(x_n)_{n\geq 0}$ and $(\varphi(x_n))_{n\geq 0}$

$$x_{n+1}=x_n-\tfrac{f(x_n)}{[x_n,x_n;f]}-\tfrac{[x_n,x_n,\varphi(x_n);f]}{[x_n,x_n;f][x_n,\varphi(x_n);f]^2}f^2(x_n) \tag{23}$$

$n=1,2,\ldots$, or, equivalently,

$$x_{n+1}=\varphi(x_n)-\tfrac{f(\varphi(x_n))}{[x_n,\varphi(x_n);f]}-\tfrac{[x_n,x_n,\varphi(x_n);f]}{[x_n,x_n;f][x_n,\varphi(x_n);f]^2}f(x_n)f\left(\varphi(x_n)\right),\quad n=1,2,\ldots. \tag{24}$$

In order to study the convergence of methods (21) or (23) we need to determine the third order derivative of f^{-1}.

By (3) for $k=3$, we get:

$$[f^{-1}(y)]'''=\tfrac{3[f''(x)]^2-f'''(x)f'(x)}{[f'(x)]^5},\quad y=f(x).$$

Regarding methods (21) or (22) the following result holds.

Theorem 4. *If the initial approximation x_1 and functions f and φ obey*

- i$_4$. *equation* (1) *has a root* $\bar{x}\in]c,d[$;
- ii$_4$. *function f is derivable up to the order* 3 *on* $[c,d]$;
- iii$_4$. $f'(x)\neq 0$, $\forall x\in[c,d]$;
- iv$_4$. *there exists $r>0$ such that* $\Delta=[\bar{x}-r,\bar{x}+r]\subseteq[c,d]$;
- v$_4$. $a=\max_{x\in\Delta}|f'(x)|$;
- vi$_4$. *function φ verifies* (18), *where* $0<b<1$;
- vii$_4$. $\rho_1=ab[\frac{ma}{6}]^{1/2}\mid\bar{x}-x_1\mid<1$, *where*

$$m=\max_{x\in\Delta}\left|\tfrac{3[f''(x)]^2-f'(x)f'''(x)}{[f'(x)]^5}\right|,$$

then the elements of the sequence $(x_n)_{n\geq 0}$ generated by (21) *remain in Δ and, moreover, the following relations hold:*

- j$_4$) $\lim x_n=\bar{x}$;
- jj$_4$) $|\bar{x}-x_{n+1}|\leq\frac{1}{ab}\sqrt{\frac{6}{ma}}\rho_1^{3^n}, n=1,2,\ldots,$ *i.e., the sequence $(x_n)_{n\geq 1}$ has the r-convergence order* 3.

Proof. We assume first that hypothesis iii$_4$ implies that function $f:[c,d]\longrightarrow F$ admits an inverse $f^{-1}:F\longrightarrow[c,d]$ and also that the root $\bar{x}$ of equation (1) is unique in the interval $[c,d]$. Conditions ii$_4$ and iii$_4$ ensure that function f^{-1} is derivable up to the order 3 on F.

By (18) and hypothesis vi$_4$ we have that for any $x_1\in\Delta$, $\varphi(x_1)\in\Delta$.

The element x_2 is given by

$$x_2=H(f(x_1),1;f(\varphi(x_1)),2;f^{-1}\mid 0)$$

and by (5) we have:

$$\bar{x}=x_2+(-1)^3[0,f(x_1),f(\varphi(x_1)),f(\varphi(x_1));f^{-1}]f(x_1)f^2(\varphi(x_1)).$$

From this relation, taking into account the mean value formula for divided differences, and also hypotheses v_4–vii_4, we get that

$$|\bar{x} - x_2| \leq \tfrac{ma^3b^2}{6} |\bar{x} - x_1|^3 . \tag{25}$$

Relation $\rho_1 < 1$ and condition (25) imply that $|\bar{x} - x_2| < |\bar{x} - x_1|$, i.e., $x_2 \in \Delta$.

In general, assuming that an element x_k of the sequence $(x_n)_{n\geq 1}$ belongs to the interval Δ, obviously $\varphi(x_n) \in \Delta$, and, similarly as above, we can prove that

$$|x_{n+1} - \bar{x}| \leq \tfrac{Ma^3b^2}{6} |\bar{x} - x_n|^3 , \quad k = 1, 2, \dots ,$$

The above relations imply jj_4. and then obviously j_4. □

Regarding the sequence generated by (23), resp. (24), the following result holds:

THEOREM 5. *If the initial approximation x_1, and the function f, φ obey*

i_5. *equation* (1) *has a root* $\bar{x} \in]c, d[$;
ii_5 *function f admits derivatives of orders up to* 3 *on* $[c, d]$;
iii_5 $f'(x) \neq 0$ *for all* $x \in [c, d]$;
iv_5 *there exists* $r > 0$ *such that* $\Delta = [\bar{x} - r, \bar{x} + r] \subseteq [c, d]$;
v_5. $a = \max\{|f'(x)| : x \in \Delta\}$;
vi_5 *function φ verifies* (18) *with* $0 < b < 1$;
vii_5 $\rho_1 = a\sqrt{\tfrac{mab}{6}}\, |\bar{x} - x_1| < 1$, *where*

$$m = \max_{x\in\Delta} \left| \tfrac{3[f''(x)]^2 - f'(x)f'''(x)}{[f'(x)]^5} \right|,$$

then the sequence $(x_n)_{n\geq 1}$ generated by (23) *is convergent, and the following relations hold:*

j_5. $\lim x_n = \bar{x}$;
jj_5. $|\bar{x} - x_{n+1}| \leq \tfrac{1}{a}\sqrt{\tfrac{6}{mab}}\, q_1^{3^n}$, $n = 1, 2, \dots$.

Proof. The proof of this theorem can be done in an analogous fashion, as for Theorem 4. □

REFERENCES

[1] COSTABILE, F., GUALTIERI, I.M. and LUCERI, R., *A new iterative method for the computation of the solution of nonlinear equations*, Numer. Algorithms, **28**, pp. 87–100, 2001.

[2] FRONTINI, M., *Hermite interpolation and a new iterative method for the computation of the roots of non-linear equations*, Calcolo, **40**, pp. 109–119, 2003.

[3] GRAU, M. *An improvement to the computing of nonlinear equation solutions*, Numer. Algorithms, **34**, pp. 1–12, 2003.

[4] OSTROWSKI, A., *Solution of Equations in Euclidean and Banach Spaces,* Academic Press, New York and London, 1973.

[5] PĂVĂLOIU, I., *Optimal efficiency index for iterative methods of interpolatory type*, Computer Science Journal of Moldova, **1**, 5, pp. 20–43, 1997.

[6] PĂVĂLOIU, I., *Approximation of the roots of equations by Aitken-Steffensen-type monotonic wequences*, Calcolo, **32**, 1–2, pp. 69–82, 1995.

[7] Păvăloiu, I., *Optimal problems concerning interpolation methods of solution of equations,* Publications de L'Institut Mathematique, **52** (**66**), pp. 113–126, 1992.

[8] Păvăloiu, I., *Optimal effiency index of a class of Hermite iterative methods, with two steps*, Rev. Anal. Numér. Théor. Approx., **29**, 1, pp. 83–89, 2000.

[9] Păvăloiu, I., *Local convergence of general Steffensen type methods*, Rev. Anal. Numér. Théor. Approx., **33**, 1, pp. 79–86, 2004.

[10] Traub, J.F., *Iterative Methods for Solutions of Equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1964.

[11] Turowicz, B.A., *Sur les derivées d'ordre superieur d'une function inverse*, Ann. Polon. Math., **8**, pp. 265–269, 1960.

Received by the editors: March 11, 2006.